Linear Algebra

B.Sc. Math honours course

Dr. Arnab Chakraborty

Assistant Professor
Indian Statistical Institute, Kolkata

Formerly

Guest Faculty
Ramakrishna Mission Vidyamandira, Belur

◇

Guest Faculty
Ramakrishna Vivekananda University, Belur

◇

Associate Professor
St. Xavier's College, Kolkata

Sarat Book Distributors
18-B, Shyamacharan De Street
Kolkata 700 073

উৎসর্গ

সম্রাট মহম্মদ বিন তুঘলকের স্মৃতির উদ্দেশ্যে
যিনি জীবদ্দশায় তাঁর উৎকট খামখেয়ালীপনায় প্রজাবর্গকে উৎপীড়ন করতেন
এবং
মৃত্যুর বহু শতাব্দী পর আজও
যাঁর বিদেহী আত্মা সিলেবাসকর্তাদের ঘাড়ে চেপে ছাত্রছাত্রীদের জ্বালিয়ে মারে।

সূচী

ছাত্রছাত্রীদের প্রতি

বিস্তর লড়াই করে এ বইয়ের তৃতীয় সংস্করণ বার করা গেল। আগের সংস্করণের আমূল পরিবর্তন করা হয়েছে এই সংস্করণে। এ বইটা আগের প্রায় দেড়গুণ সাইজের। একে তো এর মধ্যে সিলেবাস দুইবার ডিগবাজি খেয়েছে (প্রথমে CBCS এবং তারপরে NEP), তাই অনেক নতুন বিষয়ের আলোচনা ঢোকাতে হয়েছে। তার উপর আবার পুরোনো সিলেবাসের কিছু অত্যন্ত প্রাসঙ্গিক জিনিস বাদ দেওয়া হয়েছে। কিন্তু সেগুলো linear algebra শেখার পক্ষে আমি খুবই দরকারী বলে মনে করি, তাই সেগুলোও সংক্ষেপে আলোচনা করতে ছাড়িনি। যারা খালি পরীক্ষাপাশ ছাড়া আর কিছুই চাও না, তাদের সুবিধার জন্য এইসব আলোচনার শুরুতেই সাবধানবাণী লট্‌কে দেওয়া আছে।
এত সব কিছু গুছিয়ে ছাপতে পাঠাতে এত সময় লেগে গেছে যে, অনেকে আমার উপর বেজায় চটে গেছেন শুনেছি। তাঁদের কাছে আমি ক্ষমাপ্রার্থী।
আমাদের "বাংলায় বোঝানো ইংরাজি বই" সিরিজের অন্য সব বইয়ের মত এখানেও আমাদের মূল লক্ষ্য কথা যারা জানো না, তাদের জন্য বলি যে, এই সিরিজটা লেখা হচ্ছে সেই সব ছাত্রছাত্রীদের কথা মাথায় রেখে, যারা ইংরাজী মাধ্যমে পড়ছ, কিন্তু ইংরাজীর চেয়ে যাদের বাংলা বুঝতে বেশী সুবিধা হয়। এই বইগুলোতে যাবতীয় technical term ইত্যাদি ইংরাজীতেই লেখা, উত্তরগুলোও তাই। শুধু বোঝানো অংশটুকু বাংলায়।

যাদের ইংরাজী বুঝতে অসুবিধা হয় না, তারাও অনেক সময়ে বুঝতে পারো না যে একটা অংকের উত্তর কি করে ইংরাজীতে লিখতে হয়। অংক লেখার ইংরাজী আর সাধারণ ইংরাজী ঠিক এক জিনিস নয়। কি করে একটা উত্তর ইংরাজীতে গুছিয়ে লিখতে হয় সেটা এই বইতে ✎handwriting-এর মত লিখে দেখানো আছে।

Linear algebra অংকের দুনিয়ায় সবচেয়ে গুরুত্বপূর্ণ বিষয়গুলির একটা। কত জায়গায় যে এর প্রয়োগ তা যখন এই বই থেকে জানতে পারবে তখন তাজ্জব হয়ে যাবে। যারা পরীক্ষার সংকীর্ণ পরিসরের বাইরে তাকাতে চাও না, তারাও দেখবে যে যেখানে যত অংকের পরীক্ষা তাদের অধিকাংশের ক্ষেত্রেই linear algebra থেকে এক রাশ প্রশ্ন আসে। Linear algebra উঁকি মারে differential equation-এর মধ্যে, mechanics-এর আড়ালে, ইঞ্জিনিয়ারিংএর অংকে, numerical analysis-এর প্রয়োগে। তার উপরে আবার জিনিসটা বেশ মজারও বটে।

সিলেবাসের ঘূর্ণাবর্তে past years' paper-এর গুরুত্ব খানিকটা কমে গেছে। তাও নতুন সিলেবাসের যে কটা প্রশ্নপত্র হাতে পেয়েছি, সেগুলোর বিশদ আলোচনা করেছি। তাছাড়াও আছে আরো বহু প্রশ্ন, এবং তাদের বিস্তারিত উত্তর। প্রশ্নগুলো জোগাড় করতে খুব সাহায্য করেছেন আনন্দ বিশ্বাস, দীপঙ্কর দে এবং প্রতাপ মণ্ডল। আরো বহু পাঠক আগের সংস্করণগুলোর ভুলত্রুটি ধরে খুব উপকার করেছেন।
এ বইয়ের একটা webpage আছে--

`https://arnabc74.github.io/linalg`

এই বইতে কোনো ভুল-টুল পেলে সে কথাও জানিয়ে রাখব ওইখানেই। আমার Youtube channel-এও linear algebra-র টুকিটাকি বহু জিনিস আছে। বিশেষ করে "দুর্গেশগড়ের গুপ্তধন" জাতীয় সিনেমার পিছনে কীরকম ভাবে linear algebra লাগানো হয়, সে নিয়ে একটা প্লেলিস্ট। এই channel-টার হদিশ ওই webpage-এই আছে।
তোমার যদি কোনো প্রশ্ন বা মন্তব্য থাকে সেসব email (arnabc74@gmail.com) বা Whatsapp (9231542600) মারফত জানাতে পারো। তবে আমি কিন্তু বাপু প্রাইভেটে পড়াই না।

অর্ণব চক্রবর্তী

Chapter I
Linear Algebra

DAY 1
জিনিসটা কী?

এই বইতে আমরা অংকের যে শাখার কথা শিখব তার নাম linear algebra. হায়ার সেকণ্ডারীর সময়ে আমরা অল্প অল্প matrix-এর কথা শিখি। মনে করতে পারো linear algebra হল তারই পরবর্তী ধাপ। অংকের এই শাখাটার তিনটে গুরুত্বপূর্ণ বৈশিষ্ট্য আছে--

- এক, এর মত বিচিত্র ইতিহাস অংকের অন্য কোনো শাখার নেই বললেই চলে।
- দুই, এর প্রচুর প্রচুর প্রয়োগ, সুতরাং "খামোখা এসব শিখছি কেন?" প্রশ্নটা জাগলেই চট্‌জল্‌দি উত্তর পাওয়া যায়।
- তিন নম্বর বৈশিষ্ট্যটা সবচেয়ে অদ্ভুত। অংকের অন্যান্য শাখার মতই linear algebra বিষয়টাও বিভিন্ন গণিতজ্ঞের অবদানে গড়ে উঠেছে। কিন্তু তাঁরা যেভাবে বিষয়টাকে গড়ে তুলেছিলেন, সেভাবে আমরা আধুনিক যুগে মোটেই শিখি না। বিষয়টাকে ছাত্রদের মুখ চেয়ে সম্পূর্ণ ঢেলে সাজানো হয়েছে। এতে জিনিসটা শেখা অনেক সহজ হয়েছে বটে, কিন্তু "কেন লোকে প্রথম এটা ভেবেছিল?"-জাতীয় প্রশ্নের উত্তর পাওয়া কঠিন হয়ে পড়েছে।

এই তিনটে বৈশিষ্ট্য নিয়ে একটু সংক্ষিপ্ত আলোচনা না করে নিলে linear algebra-র পুরো মজাটা পাওয়া কঠিন।

1.1 প্রথম বৈশিষ্ট্য--বিচিত্র ইতিহাস

তোমরা যারা আমার মত ইতিহাসের নাম শুনলেই ভয় পাও, তাদের অভয় দিয়ে রাখি যে, ইতিহাসের আলোচনাটা খালি মজা পাবার জন্যই। আসলে linear algebra-র ইতিহাসের এমন আশ্চর্যরকমের বিশাল ব্যাপ্তি, যে সেটার উল্লেখ না করলে গল্পটা শুরুই করা যায় না! একটা উদাহরণ দিই।

Example 1: ধান, গম আর যব, এই তিনরকম শস্যের আঁটি করা হচ্ছে। একেকটা শস্যের ক্ষেত্রে আঁটির একেকরকম

সাইজ। হয়তো ধানের আঁটির বেলায় প্রতি আঁটিতে তিনটে করে ধানের চারা, গমের বেলায় হয়তো পাঁচটা করে, এইরকম। কোন্ শস্যের আঁটির সাইজ কত, তা তুমি জানো না। সেটাই তোমায় বার করতে হবে। খালি বলা আছে যে--

> 3 আঁটি ধান, 2 আঁটি গম আর 1 আঁটি যব মিলে মোট 39-খানা চারা হয়। যদি ধান, গম, যব নিতে $2, 3, 1$ আঁটি করে, তবে পেতে 34-টা চারা। আর যদি নিতে $1, 2, 3$ আঁটি করে, তবে মোট চারার সংখ্যা হত 26. বলতে পারো তিনটে শস্যের আঁটির সাইজ কী কী?

SOLUTION:এই অংকটা দেখে তুমি হয়তো ঠোঁট উল্টে বলবে--এটা আর শক্ত কি, তিনটে শস্যের জন্য আঁটির সাইজগুলো নেব x, y, z. তাহলে 3 আঁটি ধান মানে $3x$ খানা চারা, 2 আঁটি গম হল $2y$ খানা চারা, আর 1 আঁটি যবে আছে z-টা চারা। সুতরাং পাচ্ছি $3x + 2y + z = 39$. এইভাবে তিনটে শর্ত থেকে তিনটে equation পেয়ে যাব--

$$\begin{aligned} 3x + 2y + z &= 39 \\ 2x + 3y + z &= 34 \\ x + 2y + 3z &= 26. \end{aligned}$$

এবার এই তিনটে equation থেকে তিনটে unknown-কে solve করলেই হবে। এ তো আমরা অনেক আগেই শিখেছি। ■

কত আগে? যখন স্কুলে পড়তে? এবার তবে আসল কথাটা বলেই ফেলি। এই অংকটা পাওয়া গেছে চীনদেশের একটা বইতে[1], যেটা লেখা হয়েছিল খ্রীষ্টজন্মের 200 বছর আগে!!! বলা যায় যে, সেই হল linear algebra-র শুরু।
এবার আরেকটা অংক দিই।

Example 2: ধরো A, B হল দুটো square matrix. তবে বল তো $(A+B)^2$ কী করে বার করবে? তুমি বলবে এটাও সোজা অংক এখানে তো $(a+b)^2 = a^2+2ab+b^2$-এর মতই করা যাবে, খালি matrix-এর বেলায় AB আর BA সমান নাও হতে পারে, তাই এখানে হবে

$$(A+B)^2 = (A+B)(A+B) = A^2 + AB + BA + B^2.$$

■

তুমি যেভাবে এই অংকটা করলে সেই কায়দাটা প্রথম লক্ষ করে Eisenstein নামে এক পণ্ডিত ভারি উৎসাহিত হয়ে লিখেছিলেন--

> একটা matrix-এর মধ্যে অনেকগুলো সংখ্যা থাকলে কী হয়, সবটা মিলে ঠিক যেন একটা সংখ্যার মতই আচরণ করে। ওদেরকে দিব্যি সাধারণ সংখ্যাদের মতই যোগ-বিয়োগ-গুণ করা যায়, খালি গুণের বেলায় $AB = BA$ নাও হতে পারে।

এই লেখাটা কিন্তু নিতান্ত আধুনিক, 1844 খ্রীষ্টাব্দে!!! তবেই বোঝো, খ্রীষ্টপূর্ব 200 থেকে 1844 খ্রীষ্টাব্দ, এই বিশাল কালসীমা ধরে কত গবেষণার ফসল হল এই linear algebra. এবং তার পরেও আজ পর্যন্ত linear algebra নিয়ে কত গবেষণা হয়েই চলেছে।

1.2 দ্বিতীয় বৈশিষ্ট্য--প্রচুর প্রয়োগ

Linear algebra-র আরেকটা বড় বৈশিষ্ট্য হল, এটা যে কত রকম অদ্ভুত জায়গায় কাজে লাগে তার আর ইয়ত্তা নেই।

- চীনারা শস্যের হিসেব রাখার কাজ করতে গিয়ে কী করে linear algebra শুরু করেছিল তা তো বললামই।
- Einstein-এর relativity-র গবেষণা গুছিয়ে লিখতেও linear algebra বহুল ব্যবহার হয়েছে।
- Gauss (গাউস) নামের একজন জার্মান গণিতজ্ঞ linear algebra দিয়ে একটা astronomy-র সমস্যার চমৎকার সমাধান বার করেন। সৌরজগতে যেমন গ্রহগুলো আছে, তেমনি আছে অনেকগুলো বিশাল বিশাল এবড়ো-খেবড়ো পাথরের চাঁই, যাদের বলে asteroid. এরাও নির্দিষ্ট পথে সূর্যের চারদিকে পাক খায়। এদের মধ্যে দুটোর গতিপথ নির্ধারণ করেছিলেন Gauss এবং এই কাজের জন্য তিনি linear algebra-র একটা কায়দা বার করেছিলেন, যেটা Gaussian elimination নামে বিখ্যাত। মজার কথা হল সেই যে চীনা বইটার কথা বলেছিলাম, সেইখানেও ঠিক একই কায়দার কথা অত দিন আগেই বলা ছিল, যদিও Gauss সম্ভবতঃ সেটা জানতেন না।
- তোমরা parabola, ellipse ইত্যাদির কথা হায়ার সেকণ্ডারীতে পড়েছ। সেখানে ওদের axis-গুলোকে আমরা x-axis বা y-axis-এর সঙ্গে parallel (সমান্তরাল) নিতাম (Fig 1)। কিন্তু বাস্তবে অনেক সময়েই হেলানো axis নিয়েও কাজ করতে হয় (Fig 2)। তখন তাদের equation-গুলো এত জটিল দেখায় যে, তা থেকে বোঝা শক্ত সেটা parabola নাকি ellipse নাকি অন্যকিছু। এই সনাক্ত করার জন্যেও linear algebra কাজে দেয়! এটাও বার করেছিলেন সেই Gauss.
- আজকাল Google-এর কথা কে না জানে! তুমি কোনো একটা জিনিস নিয়ে তথ্য চাইলে Google ঠিক দরকারী website-গুলো খুঁজে বার করে দেয়। কী করে পারে? শুনলে আশ্চর্য হবে যে, এখানেও linear algebra ব্যবহার হয়! আমরা এই বইতে সেটাও শিখব।

[1] বইটার নাম বাংলা করলে দাঁড়ায় "অংকের কারিকুরি নিয়ে নয়টি অধ্যায়"

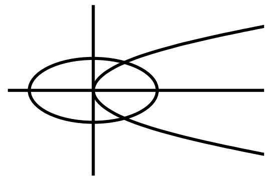


Fig 1

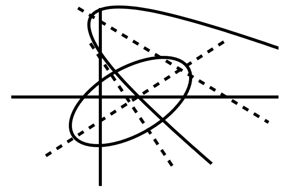


Fig 2

- Robot-রা যে নড়াচড়া করে বা বিভিন্ন 3D কম্পিউটার গেমের চরিত্রগুলো যে চলেফিরে বেড়ায়, সেটাও সম্ভব হয় linear algebra-র দৌলতেই।

1.3 তৃতীয় কথা--ভূতের রাজা দিল বর

এইবারে শোন্, মোদের কেমনে হয়েছে এই হাল...
--হীরক রাজার দেশে

এত বেশী প্রয়োগ অংকের আর কোনো শাখায় আছে কিনা সন্দেহ। কিন্তু তাতে একটা সমস্যা হয়েছে। বিভিন্ন মানুষ বিভিন্ন সময়ে বিভিন্ন দিক থেকে linear algebra-কে বার বার করে আবিষ্কার করেছে। ভাবতেও পারে নি যে, আগে লোকে একই জিনিস অন্য সূত্রেও বার করে বসে আছে। যেমন Gauss কি ঘুণাক্ষরেও সন্দেহ করেছিলেন যে, astronomy-র জন্য তিনি যে কায়দাটা সদ্য বার করলেন, সেটা বহু শতাব্দী আগে চীন দেশে কৃষিকাজের সমস্যা সমাধানের জন্য কেউ বার করে বসে আছে? এইভাবে বহুদিন পর্যন্ত linear algebra বলে অংকের কোনো শাখা তৈরীই হয় নি। আজ আমরা যেগুলো linear algebra-র সম্পদ বলে জানি, সেসব এখানে সেখানে ছড়িয়ে ছিল। যখন Einstein তাঁর রিলেটিভিটির গবেষণা প্রকাশ করলেন, তিনি এই সব ধারণার অনেকগুলো ব্যবহার করেছিলেন। কিন্তু তাঁর notation এত ভয়ানক রকমের জটিল ছিল যে, তাঁর গবেষণা লোকের মাথায় ঢুকতে বেশ সময় লেগেছিল। কিন্তু ধারণাটার গুরুত্ব বোঝার পর পণ্ডিতেরা উঠে পড়ে লাগলেন পুরো জিনিসটাকে গুছিয়ে লেখার জন্য। এবং সেই প্রথম linear algebra নামে একটা নতুন শাখা আত্মপ্রকাশ করল। সবকিছু সাজিয়ে লিখতে গিয়ে দেখা গেল যে, এত দিন যেভাবে বিষয়টা ছন্নছাড়ার মত গড়ে উঠেছিল, সেভাবে না এগিয়ে অন্যভাবে এগোলে সবকিছু অনেক সহজে বোঝা যায়, মনে রাখতেও সুবিধা হয়। তাই আজ linear algebra-র বইপত্র সব সেই কায়দাতেই লেখা হয়, এবং সেই কারণেই এই সুদীর্ঘ গবেষণার ফসল ছাত্ররা সহজেই শিখতে পারে। এর ফলে বিভিন্ন সংজ্ঞার পিছনের ধারণাগুলো একেবারে বেমালুম বদলে যায়, যেন ভূতের রাজার বরে! তবে এইভাবে শেখার খালি একটাই মুস্কিল, এটা একটা কৃত্রিম পদ্ধতি। একটা অংক কীভাবে করতে হবে সেটা এতে চমৎকার বোঝা যায়, কিন্তু কেন করতে হবে, বা মানুষ আদৌ কেন এই ধারণাটা ভেবেছিল তার হদিশ পাওয়া যায় না। যেমন, তোমরা হায়ার সেকণ্ডারীতে matrix-এর কথা শিখেছ, এবং determinant-এর কথাও শিখেছ। হয়তো প্রশ্ন জেগেছে--খামোখা একটা matrix-এর determinant লোকে বানালো কেন? আসলে কিন্তু লোকে মোটেই কখনো একটা matrix নিয়ে তার determinant-এর সংজ্ঞা বানাতে বসে নি। Determinant-এর ধারণা এসেছে matrix-এর বহু আগে। এখনকার দৃষ্টিভঙ্গী থেকে এ কথাটা শুনতেও অবাক লাগে--matrix-ই যদি না থাকে, তবে তার determinant-এর প্রশ্নটা এল কী করে?

ব্রহ্মাণ্ড ছিল না যখন মুণ্ডমালা কোথা পেলি?
--রামপ্রসাদ

এভাবে গোড়াতেই নানা রকম সঙ্গত প্রশ্নের সদুত্তর না পাওয়ায় অনেক ছাত্রই linear algebra পড়তে গিয়ে খানিকটা মুষড়ে পড়ে। আমরা তাই মূলতঃ আধুনিক কায়দাই অনুসরণ করলেও মাঝে মাঝে একটু ইতিহাসের গল্প করে নেব যাতে চিন্তার খেইটা হারিয়ে না যায়।

1.4 কী কী জানা আছে

হায়ার সেকণ্ডারীতে যেটুকু matrix শিখেছিলে, সেগুলো এখানে কাজে লাগবে। একটু মনে করে নেওয়া যাক।

একটা **matrix** মানে কী, তা তো জানোই--কিছু সংখ্যা যেগুলোকে একটা rectangle-এর মত সাজিয়ে লেখা হয়েছে, যেমন--

$$\begin{bmatrix} 1 & -\frac{1}{2} & 0 & \pi \\ 2 & 3 & 100 & 1 \\ 0 & 0 & 6 & 4 \end{bmatrix}.$$

প্রত্যেকটা horizontal সারিকে বলে এককেটা **row**, আর vertical সারিগুলোকে বলে এককেটা **column.** এখানে matrix-টার 3-টে row আর 4-টে column. তাই একে বলব একটা 3×4 matrix. লক্ষ কর যে প্রথমে লিখেছি row-এর সংখ্যা, তার পরে column-এর সংখ্যা। Matrix-এর ভিতরের প্রত্যেকটা সংখ্যাকে বলে এককেটা **entry.** এখানে একটা entry হল 100. এটা আছে দ্বিতীয় row আর তৃতীয় column-এ। তাই একে আমরা বলব $(2,3)$-th entry. এখানেও দেখো আগে row, পরে column লিখলাম। যদি গোটা matrix-টার নাম দিই A, তবে এই entry-টাকে লিখব a_{23} বা $a_{2,3}$. এটাই প্রচলিত নিয়ম--একটা matrix-এর নাম সাধারণতঃ কোনো বড় হাতের ইংরাজি অক্ষর দিয়ে দেওয়া হয়, আর entry-গুলোকে বোঝানো হয় ছোটো হাতে সেই একই অক্ষর দিয়ে। এখানে a-র নীচে ওই যে ছোটো করে 2 আর 3 লিখেছি, ওদের বলে **subscript** বা **index**.

একটা matrix-এ যদি খালি একটাই column থাকে, তবে তাকে অনেক সময়ে **column vector** বলে। তেমনি যদি খালি একটাই row থাকে, তবে বলে **row vector.** যেমন $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ হল একটা column vector, আর $[1 \quad 2 \quad 3]$ হল একটা row vector. আর যদি আলাদা করে row vector বা column vector না বলে খালি vector বলি, তবে সাধারণতঃ column vector-ই বোঝায়।

যদি একটা matrix-এ row আর column-এর সংখ্যা সমান হয়, তবে তাকে বলে **square matrix**. কেন বলে তা তো বুঝতেই পারছ। যদি এই square-টার দুটো diagonal (কর্ণ) আঁকো, তবে একটা যাবে উত্তর-পশ্চিম থেকে দক্ষিণ-পূর্বে আর অন্যটা দক্ষিণ-পশ্চিম থেকে উত্তর-পূর্বে (Fig 3)। এদের মধ্যে আমরা সাধারণতঃ প্রথমটাকে নিয়েই মাথা ঘামিয়ে থাকি, ওকে বলে **leading diagonal.** অন্য diagonal-টার কোনো আলাদা নাম পর্যন্ত নেই, কারণ ওটা কখনোই ব্যবহার হয় না। তাই কখনো যদি একটা matrix-এর diagonal-এর কথা শোনো, তবে নিশ্চিত জানবে যে leading diagonal-এর কথাই হচ্ছে। এই diagonal বরাবর যে সব entry আছে তাদের বলে **diagonal entry**. প্রথমটা হল $(1,1)$-th entry, তার পরেরটা $(2,2)$-th, এইরকম, মানে যত নম্বর row তার ঠিক তত নম্বর column-এর entry. সাধারণতঃ matrix-টা square না হলে তার diagonal নিয়ে লোকে বড় একটা মাথা ঘামায় না। তবে যদি দেখো যে A একটা matrix যেটা square নয়, কিন্তু তাও লোকে ওর diagonal entry-দের নিয়ে আলোচনা করছে, তবে জানবে যে এখানেও সেই একই নিয়ম অনুসরণ করা হচ্ছে--যত নম্বর row তার তত নম্বর entry. মানে diagonal entry-রা এখানেও $a_{11}, a_{22}, \ldots$ ইত্যাদি। যদিও Fig 4 দেখলেই বুঝবে এরা মোটেই জ্যামিতিক অর্থে diagonal বরাবর নেই। Diagonal entry ছাড়া বাকি entry-দের বলে **off-diagonal** entry.

দুটো matrix-কে কী করে যোগ বা বিয়োগ করতে হয়, তাও তোমার অজানা নয়। এর জন্য দুটো matrix-কেই একই সাইজের হতে হবে। যোগ-বিয়োগ করা হয় entry ধরে ধরে। প্রথম matrix-টার $(1,1)$-th entry-র সঙ্গে দ্বিতীয় matrix-এর $(1,1)$-th entry, এইরকম।

Fig 3

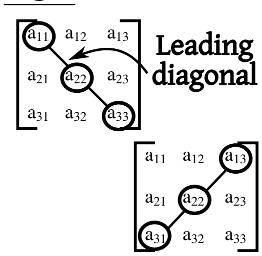
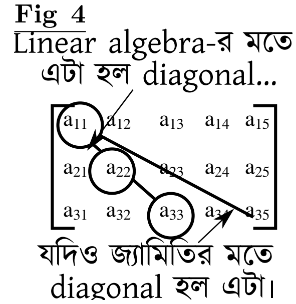


Fig 4



Example 3: $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ আর $B = \begin{bmatrix} 10 & 12 & 33 \\ -4 & 53 & 26 \end{bmatrix}$ হলে $A + B$ বার করো।

SOLUTION:

$$A + B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 10 & 12 & 33 \\ -4 & 53 & 26 \end{bmatrix} = \begin{bmatrix} 1+10 & 2+12 & 3+33 \\ 4-4 & 5+53 & 6+26 \end{bmatrix} = \begin{bmatrix} 11 & 14 & 36 \\ 0 & 58 & 32 \end{bmatrix}.$$

■

যোগ-বিয়োগটা বেশ সহজ জিনিস। আরেকটা সহজ জিনিস হল **transpose** করা, অর্থাৎ matrix-টার row-গুলোকে column (এবং তার ফলে column-গুলোকে row) করে দেওয়া, যেমন

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

হলে তার transpose হবে

$$A' = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

এই বইতে আমরা কোনো matrix A-র transpose-কে A' লিখব। অনেকে A^T বা A^t-ও লেখে। যদি A-র entry-গুলো complex number হয়, তাহলে সাধারণতঃ লোকে খালি transpose না নিয়ে **conjugate transpose** বার করে, মানে transpose করার পর প্রত্যেকটা entry-র **conjugate** বার করে। Conjugate কাকে বলে মনে আছে তো? যদি একটা complex সংখ্যা নাও $a + ib$, যেখানে $a, b \in \mathbb{R}$, তবে তার conjugate হবে $a - ib$. লেখার সময়ে conjugate বোঝায় মাথার উপর দাগ টেনে, যেমন

$$\overline{2 + 3i} = 2 - 3i.$$

সুতরাং একটা matrix A-র transpose-এর conjugate নেওয়া মানে হল $\overline{A'}$ বার করা। একে সংক্ষেপে লেখে A^*, আর মুখে বলে A-star বা conjugate transpose of A.

Example 4: যদি

$$A = \begin{bmatrix} 1+9i & 4i \\ 2 & 5 \\ 3-i & 6 \end{bmatrix}$$

হয়, তবে A^* বার কর।

SOLUTION: প্রথমে A' বার করি--

$$A' = \begin{bmatrix} 1+9i & 2 & 3-i \\ 4i & 5 & 6 \end{bmatrix}.$$

এবার প্রত্যেকটা entry-র conjugate করি--

$$A^* = \begin{bmatrix} 1-9i & 2 & 3+i \\ -4i & 5 & 6 \end{bmatrix}.$$

লক্ষ কর যে conjugate নেওয়ায় real entry-গুলোর কোনো পরিবর্তন হল না। সুতরাং বুঝতেই পারছ যে, যদি A একটা real matrix হয় (অর্থাৎ প্রত্যেকটা entry-ই real সংখ্যা হয়), তবে A^* আর A' একই জিনিস। ■

Exercise 1: A কোন্ ধরণের matrix হলে $A + A'$ বার করা সম্ভব? ■

Exercise 2: যদি A আর B দুটো 5×5 সাইজের complex matrix হয়, তবে নীচের বক্তব্যগুলোর মধ্যে কোন্টা কোন্টা সঠিক?

(A) $A - A'$-এর সবগুলো entry-র যোগফল শূন্য।

(B) $(A + B)^* = A^* + B^*$.

(C) যদি $C = B - B^*$ হয়, তবে C-এর diagonal entry-গুলো সব শূন্য।

(D) $(A')' = A$.

■

Exercise 3: এমন একটা 2×2 সাইজের complex matrix দিতে পারো, A, যাতে $A + A^*$-এর $(1, 1)$-th entry হয় i? ■

DAY 2
Matrix multiplication

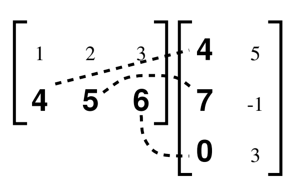
দুটো matrix-কে কী করে গুণ করতে হয়, সেটা হায়ার সেকণ্ডারীতেই শেখানো হয়ে থাকে। তবে দুটো matrix দিয়ে যদি গুণ করতে বলি, তবে অনেকেরই দেখেছি হাত একটু কাঁপে। তাই গোড়া থেকেই শুরু করা যাক। যোগ-বিয়োগ বা transpose করার চেয়ে গুণটা নিঃসন্দেহে একটু কঠিন। ধরো দুটো matrix আছে A আর B. তাহলে আমরা AB বার করতে পারব যদি A-র column-এর সংখ্যা আর B-এর row-এর সংখ্যা সমান হয়। যদি A হয় $m \times n$ আর B হয় $n \times p$, তবে AB হবে একটা $m \times p$ আকারের matrix. গুণটা কী করে করতে হয়, সেটা একটা উদাহরণ দিয়ে মনে করে নেওয়া যাক।

Example 5: ধরো

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{ আর } B = \begin{bmatrix} 4 & 5 \\ 7 & -1 \\ 0 & 3 \end{bmatrix}.$$

তবে AB বার করো।

Fig 5



SOLUTION: এই উদাহরণে A হল 2×3 আর B হল 3×2. এখানে A-র column-সংখ্যা আর B-এর row-সংখ্যা সমান (দুটোই 3). সুতরাং AB বার করা যাবে, এবং গুণফলটার সাইজ হবে 2×2. তার মানে উত্তরটায় চারটে entry থাকবে। আমরা উদাহরণস্বরূপ একটা entry বার করার কায়দাটা বিশদভাবে দেখাব। ধরো $(2,1)$-th entry-টা। এর জন্য A-র 2 নম্বর row আর B-এর 1 নম্বর column-টা দেখো। বোঝার সুবিধার জন্য Fig 5-এ এদের মোটা করে দেখিয়েছি। এবার row-টার প্রথম সংখ্যার সাথে column-টার প্রথম সংখ্যা গুণ কর, তারপর একই ভাবে দ্বিতীয় সংখ্যাদুটোকেও গুণ কর, তারপর তৃতীয়গুলোকেও, যেমন ছবিতে লাইন টেনে দেখিয়েছি। এই তিনটে গুণফলকে যোগ করে দিলে পাবে AB-র $(2,1)$-th entry:

$$4\times 4+5\times 7+6\times 0=51.$$

একই কায়দায় বাকি entry-গুলোও বার করা যাবে। সব মিলিয়ে পাবে--

$$AB=\begin{bmatrix}18 & 12\\ 51 & 33\end{bmatrix}.$$

■

Exercise 4: $A=\begin{bmatrix}2 & 3 & 4\\ 4 & 5 & 5\end{bmatrix}$ আর $B=\begin{bmatrix}0 & 2 & 6\\ 0 & 0 & 2\end{bmatrix}$ হলে AB', $A'B$, AB এবং $(AA')^2$ বার কর (বা যদি বার করা অসম্ভব হয়, সেটা বলো)। এখানে $(AA')^2$ মানে হল $(AA')(AA')$. ■

2.1 কিন্তু কেন?

কী করে গুণ করতে হয় সেটা তো বোঝা গেল, কিন্তু এমন বিদ্‌ঘুটে একটা নিয়ম এল কোথা থেকে? এর সবচেয়ে ভালো উত্তরটা জানতে পারবে পরে, যখন আমরা linear transformation বলে একটা জিনিস শিখব। আপাততঃ তিনটে সহজ উদাহরণ বলি, যেখানে এই বিদ্‌ঘুটে নিয়মটা খুব স্বাভাবিকভাবেই আসে।

2.1.1 সুদে আসলে

ছোটোবেলায় আমরা সবাই একটু-আধটু সুদ-আসলের অংক করেছি। অংকগুলো এরকম হত--একটা ব্যবসাতে টাকা খাটালে বছরে 10% হারে সুদ পাওয়া যায়, মানে তুমি 100 টাকা খাটালে প্রতি বছর[2] 10 টাকা সুদ পাবে। যদি তুমি 500 টাকা খাটাও, তবে কত সুদ পাবে? আমরা ছোটোবেলা থেকেই জানি যে, এখানে গুণ করলেই উত্তরটা পাওয়া যায়--

সুদ $=$ কত টাকা খাটানো হয়েছে $\times$ সুদের হার।

তার মানে এই উদাহরণে সুদ হবে $500\times 0.10=50$ টাকা। এবার এর একটু কঠিন সংস্করণ দেখা যাক। সাধারণতঃ লোকে খালি একটা জায়গাতেই টাকা খাটায় না। কারণ সুদের হার ওঠানামা করে, এই মাসে যে কোম্পানীটা লাভ করল পরের মাসে সে যে ডিগবাজি খাবে না, এমন স্থিরতা নেই। তাই লোকে বিভিন্ন কোম্পানি মিলিয়ে মিশিয়ে টাকা খাটায়। কিছু লোক টাকার বেশী অংশটা সরকারী কোম্পানীতে রাখতে ভালোবাসে। সরকারী কোম্পানিগুলোর সুদের হার হয়তো বেশী নয়, কিন্তু হারটা মোটামুটি একইরকম থাকে। আবার একেকটা ভুঁইফোঁড় কোম্পানি থাকে, তারা হয়তো দুই বছর বেজায় চড়া হারে সুদ দিয়ে, পরের বছর তলিয়ে গেল। কেউ কেউ চড়া সুদের লোভে এরকম জায়গাতেও অনেক টাকা খাটায়। এরকম দুটো কোম্পানীর সুদের হিসেব দিলাম পর পর চার বছরের জন্য।

	প্রথম বছর	দ্বিতীয় বছর	তৃতীয় বছর	চতুর্থ বছর
সরকারী	0.05	0.06	0.05	0.05
বেসরকারী	0.10	0.15	0.04	0.01

এবার তিনজন লোক এই দুই কোম্পানীতে কত কত টাকা খাটিয়েছেন বলি--

[2] এখানে আমরা সরল সুদের কথা বলছি, চক্রবৃদ্ধি সুদের জটিলতায় যাচ্ছি না, যেখানে এই বছরের সুদের উপরে আবার পরের বছর সুদ পাওয়া যায়।

	সরকারী	বেসরকারী
তারক রায়	10000	1000
জয়মল দুবে	0	50000
দেবাশিস বসু	20000	20000

তারকবাবু রিটায়ার্ড স্কুলটিচার, পারলে পুরো টাকাটাই সরকারী কোম্পানিতেই রাখেন, নেহাত জামাইয়ের কথায় অল্প খানিকটা বেসরকারী কোম্পানীতেও রেখেছেন। জয়মল দুবে চট্ করে টাকা কামাবার তালে আছেন। তিনি সব টাকাই ঢেলেছেন বেসরকারী কোম্পানিতে। দেবাশিসবাবু অবশ্য দুই দিক বাঁচিয়ে চলার পক্ষপাতী।
বল তো তৃতীয় বছরে তারকবাবু মোট কত সুদ পেয়েছেন? সহজ উত্তর, সরকারী থেকে 10000×0.05, আর বেসরকারী থেকে 1000×0.04. সুতরাং মোট সুদ হল

$$10000 \times 0.05 + 1000 \times 0.04 = 540.$$

একইভাবে তুমি বার করে ফেলতে পারবে প্রত্যেকে প্রতিবছর কত করে সুদ পেয়েছেন। আমরা তার তালিকাটা নীচে দিলাম। তারকবাবুর তৃতীয় বছরের সূদটা (যেটা আমরা এক্ষুণি কষলাম), সেটা মোটা করে দেখিয়েছি।

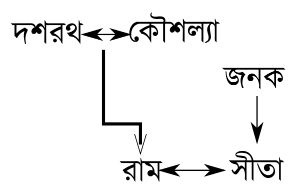
	প্রথম বছর	দ্বিতীয় বছর	তৃতীয় বছর	চতুর্থ বছর
তারক রায়	600	750	**540**	510
জয়মল দুবে	5000	7500	2000	500
দেবাশিস বসু	3000	4200	1800	1200

যেভাবে এই তালিকাটা বার করলাম, তার মধ্যে এমন কিছুই লাগে নি যা আমরা ছোটোবেলায় স্কুলেই শিখি নি। অথচ ভালো করে তাকিয়ে দেখো তো, এইটাই তো আসলে matrix multiplication–

$$\begin{bmatrix} 600 & 750 & 540 & 510 \\ 5000 & 7500 & 2000 & 500 \\ 3000 & 4200 & 1800 & 1200 \end{bmatrix}$$
$$= \begin{bmatrix} 10000 & 1000 \\ 0 & 50000 \\ 20000 & 20000 \end{bmatrix} \times \begin{bmatrix} 0.05 & 0.06 & 0.05 & 0.05 \\ 0.10 & 0.15 & 0.04 & 0.01 \end{bmatrix}$$

সুতরাং ছোটোবেলার সেই "সুদ $=$ কত টাকা খাটানো হয়েছে $\times$ সুদের হার" এখানেও প্রযোজ্য, খালি এখানে তিনটে জিনিসই হল matrix, আর ডানদিকের গুণটা হল matrix multiplication.
আর এই উদাহরণটা থেকে আরেকটা জিনিসও সহজে বোঝা যায়--কেন matrix multiplication-এর সময়ে প্রথম matrix-টার column-সংখ্যা দ্বিতীয়টার row-সংখ্যার সমান হতে হয়। এখানে টাকা খাটানোর matrix-এ column ছিল দুটো, কারণ কোম্পানির সংখ্যা দুই, এবং ঠিক সেই কারণেই সুদের হারের matrix-এর row-এর সংখ্যাও দুই।

Fig 6



2.1.2 রামায়ণের উদাহরণ

রামায়ণের গল্পে আছে দশরথ আর কৌশল্যার ছেলে রাম। রামের স্ত্রী সীতা, এবং সীতার বাবা হলেন মিথিলার রাজা জনক। Fig 6 দেখো। যদি আমরা জিজ্ঞাসা করি "কার বাবা কে?" তবে উত্তরটা একটা matrix-এর আকারে লেখা যায়--

	দশরথ	কৌশল্যা	রাম	সীতা	জনক
দশরথ	0	0	1	0	0
কৌশল্যা	0	0	0	0	0
রাম	0	0	0	0	0
সীতা	0	0	0	0	0
জনক	0	0	0	1	0

এখানে (i,j)-th entry $= 1$ মানে i-th মানুষটা j-th মানুষটার বাবা। যেমন দশরথের row-তে রামের নীচে 1 আছে, কারণ দশরথ রামের বাবা। একইভাবে যদি প্রশ্ন করি "কে কার স্ত্রী?" তার উত্তরে পাব এই matrix-টা--

	দশরথ	কৌশল্যা	রাম	সীতা	জনক
দশরথ	0	0	0	0	0
কৌশল্যা	1	0	0	0	0
রাম	0	0	0	0	0
সীতা	0	0	1	0	0
জনক	0	0	0	0	0

এবার একটা তৃতীয় প্রশ্ন করি (যেটা লোকে সাধারণতঃ রামায়ণ পড়তে গিয়ে করে না)-- 200-র $\frac{1}{2}$ কত? অবশ্য উত্তরটা সোজাই--

$$200\text{-র } \frac{1}{2} = \frac{1}{2} \times 200 = 100.$$

লক্ষ কর এখানে কীভাবে গুণ চিহ্ন এসে গেল। এবার আবার রামায়ণের প্রসঙ্গে যাই। তুমি চট্ করে "স্ত্রী" matrix-টাকে "বাবা" matrix দিয়ে গুণ করে দাও দেখি! উত্তর পাবে--

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

এইটা কিসের matrix জানো তো? এটা হল "বাবার স্ত্রী" মানে "মা"-র matrix–

	দশরথ	কৌশল্যা	রাম	সীতা	জনক
দশরথ	0	0	0	0	0
কৌশল্যা	0	0	1	0	0
রাম	0	0	0	0	0
সীতা	0	0	0	0	0
জনক	0	0	0	0	0

মিলিয়ে দেখো, এখানে কৌশল্যা রামের মা। এই matrix-টাও তাই বলছে, কারণ (কৌশল্যা, রাম) entry-টা দেখা যাচ্ছে 1. ঠিক যেমন 200-র $\frac{1}{2}$ মানে $\frac{1}{2} \times 200$, তেমনি "বাবার স্ত্রী" মানে "স্ত্রী"$\times$"বাবা"!!!
আর যদি "বাবা"-কে "স্ত্রী" দিয়ে গুণ করতে তবে পেতে "স্ত্রীর বাবা" মানে "শ্বশুর"-এর matrix, করেই দেখো!

এই উদাহরণটা থেকেই বুঝবে matrix multiplication কেন commutative হয় না। হাজার হোক, বাবার স্ত্রী আর স্ত্রীর বাবা তো এক জিনিস নয়!

2.1.3 Facebook-এর উদাহরণ

ত্রেতাযুগের রামায়ণ থেকে সোজা চলে আসি কলিযুগের Facebook-এ। উদাহরণটার মূল জিনিসটা এখানেও একইরকম, বিভিন্ন মানুষের মধ্যে সম্পর্ক। আর Facebook-এ সম্পর্ক বলতে তো প্রথমেই মনে আসে বন্ধুত্বের কথা। আমাদের উদাহরণে পাঁচটি চরিত্র--আলম, বিভাস, চঞ্চল, দীপান্বিতা আর এলা। এদের মধ্যে Facebook-এ কে কার friend সেটা Fig 7-এ লাইন টেনে টেনে দেখিয়েছি। রামায়ণের অভিজ্ঞতা থেকেই তুমি জানো এরকম সম্পর্ককে কীভাবে matrix দিয়ে লিখতে হয়--

	আলম	বিভাস	চঞ্চল	দীপান্বিতা	এলা
আলম	0	0	1	1	1
বিভাস	0	0	0	1	1
চঞ্চল	1	0	0	1	0
দীপান্বিতা	1	1	1	0	0
এলা	1	1	0	0	0

ধরি এই matrix-টার নাম F. এবার যদি F^2 বার করি, তবে সেটা কীসের matrix হবে? আন্দাজ করা শক্ত নয় যে উত্তর হল--friend of friend-এর matrix হবে--

	আলম	বিভাস	চঞ্চল	দীপান্বিতা	এলা
আলম	3	2	1	1	0
বিভাস	**2**	2	1	0	0
চঞ্চল	1	1	2	1	1
দীপান্বিতা	1	0	1	3	2
এলা	0	0	1	2	2

লক্ষ করো যে F^2-এর entry-গুলো কিন্তু খালি 0 আর 1 নয়। কয়েক জায়গায় 2 আর 3-ও এসেছে। যেমন F^2-এর (বিভাস, আলম) entry-টা হল 2 (মোটা করে দেখিয়েছি)। এর মানে কী? যারা Facebook কর, তাদের কিন্তু উত্তরটা একটু ভেবেই বলতে পারা উচিত-- এর মানে বিভাস আর আলমের 2 জন mutual friend আছে, অর্থাৎ এমন ঠিক 2 জন আছে যারা বিভাসেরও বন্ধু আবার আলমেরও বন্ধু। Fig 7-এর সঙ্গে মিলিয়ে দেখে নাও, সত্যিই বিভাস আর আলমের ঠিক দুইজনই mutual friend আছে--এলা এবং দীপান্বিতা। Facebook-এ তুমি যখন কারো page-এ যাও, তখন Facebook ঠিক এই matrix multiplication-এর কায়দাতেই বার করে তোমার সঙ্গে তার কতজন mutual friend আছে!

Exercise 5: আচ্ছা, বলো তো F^2-এর diagonal entry-গুলোর মানে কী? ■

Fig 7

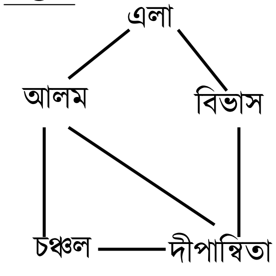
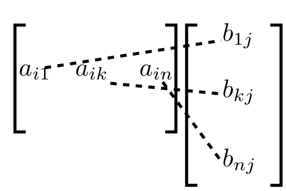


Fig 8



Fig 9



Exercise 6: এবার Fig 8-এর ভিত্তিতে matrix multiplication দিয়ে mutual friend-দের সংখ্যা বার করো তো। ■

2.2 Product-এর entry লেখার কায়দা

ধরো দুটো matrix দিলাম $A_{m\times n}$ আর $B_{n\times p}$. এদের গুণ করে পাওয়া যায়

$$C_{m\times p} = AB.$$

এই C-কে বলব A আর B-এর **product** (গুণফল)। যদি তোমাকে কেউ বলে C-এর (i,j)-th entry-টা লিখে দিতে, তবে সেটা করার একটা কায়দা আছে, যেটা অনেক সময়ে কাজে লাগে। এর জন্য A আর B-এর entry-গুলো লাগবে, এদের নাম দেব a_{ij} আর b_{ij}. এবার যদি C-এর (i,j)-th entry-টাকে আমরা c_{ij} বলি, তবে তার ফর্মুলা লেখার জন্য এইভাবে ধাপে ধাপে এগোও--

- প্রথমে পাশাপাশি a আর b লেখো--

  $$a..b..$$

  এদের প্রত্যেকের নীচে দুটো করে subscript বসবে, তাই ডট্ ডট্ দিয়ে জায়গাগুলো রেখেছি।

- এবার আমাদের কাজ হল subscript-গুলো বসানো। যেহেতু আমরা (i,j)-th entry বার করতে চাই, তাই a-র নীচে প্রথম subscript-টা নেব i, আর b-এর নীচে শেষ subscript-টা নেব j. মানে দাঁড়ালো--

  $$a_{i.}b_{.j}$$

- এখনো বাকি আছে a-র দ্বিতীয় subscript আর b-এর প্রথম subscript-টা। এর জন্য i,j ছাড়া যা খুশি কিছু অক্ষর নাও, ধরো k নিলাম। দুই জায়গাতেই k বসিয়ে দাও--

  $$a_{ik}b_{kj},$$

- এখানে k জিনিসটা আমাদের আমদানি। পুরোটার বাঁদিকে একটা summation হবে এই k-র উপরে। লক্ষ কর k হল a-র দ্বিতীয় subscript, মানে k বোঝাচ্ছে A-র column-দের। আবার k-টা b-র প্রথম subscript-ও বটে, মানে ওটা B-র row-দেরও বোঝাচ্ছে। যেহেতু AB গুণফলটা আদৌ বার করা যাচ্ছে, তার মানে A-র column-সংখ্যা আর B-র row-সংখ্যা সমান। আমাদের উদাহরণে এই সংখ্যাটা হল n. তাই summation-টা হবে $k=1$ থেকে $k=n$ পর্যন্ত। ব্যস্, সব মিলিয়ে দাঁড়ালো--

  $$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Fig 9 দেখলেই বুঝবে যে এটা কেন কাজ করছে। চট্ করে লেখার জন্য এই কায়দাটা জানা থাকা ভালো।

Exercise 7: $A_{5\times 4}$ আর $B_{4\times 10}$-কে গুণ করে পাওয়া গেছে $C=AB$. তাহলে c_{23}-কে A আর B-এর entry-দের দিয়ে লেখো। ■

যদি দুইয়ের বেশী matrix-এর গুণফল নিয়ে কাজ করতে হয়, তাহলেও এই কায়দাটা খাটে, যেমন নীচের উদাহরণে--

Example 6: ধরো চারটে matrix আছে $A_{m\times n}, B_{n\times p}, C_{p\times q}$ আর $D_{q\times r}$. তবে $ABCD$ গুণফলটার (i,j)-th entry কি হবে?

SOLUTION: প্রথমে লিখে নেব $abcd$, তারপর subscript-গুলো বসাব--

$$a..b..c..d...$$

Subscript-দের মধ্যে একেবারে বাঁদিকে i আর একেবারে ডানদিকে j বসাব--

$$a_i.b..c..d._j.$$

এবার a-র দ্বিতীয় আর b-এর প্রথম subscript যা খুশি কিছু নেব, ধরো k নিলাম--

$$a_{ik}b_k.c..d._j.$$

একইভাবে b-এর দ্বিতীয় আর c-এর প্রথম subscript-ও যা খুশি কিছু নেব ধরো ℓ. আর একইভাবে c আর d-এর জন্যও যা খুশি কিছু নেব, ধরো t. তার মানে পেলাম--

$$a_{ik}b_{k\ell}c_{\ell t}d_{tj}.$$

নতুন জিনিস এখানে এনেছি তিনটে, k, ℓ আর t. এদের প্রত্যেকের উপর summation চাই--

$$\sum_k \sum_\ell \sum_t a_{ik}b_{k\ell}c_{\ell t}d_{tj}.$$

এখানে k ব্যবহার হয়েছে $A_{m\times n}$ এবং $B_{n\times p}$-এর জন্য। এদের সাইজের মধ্যে কমন আছে n. সুতরাং k যাবে 1 থেকে n অব্ধি। একইভাবে ℓ যাবে 1 থেকে p পর্যন্ত, আর t যাবে 1 থেকে q পর্যন্ত--

$$\sum_{k=1}^n \sum_{\ell=1}^p \sum_{t=1}^q a_{ik}b_{k\ell}c_{\ell t}d_{tj}.$$

ব্যস্ এই হল উত্তর। এত কথা বোঝাতেই যা সময় লাগল। একবার কায়দাটা বুঝে গেলে ধাঁ করে লিখে ফেলা যায়। ■

Exercise 8: $A_{5\times 4}$, $B_{4\times 10}$ আর $C_{10\times 2}$-কে গুণ করে পাওয়া গেছে $D = ABC$. তাহলে d_{21}-কে A, B আর C-এর entry-দের দিয়ে লেখো। ■

2.3 দুই ধরণের গুরুত্বপূর্ণ matrix

দুই ধরণের matrix আছে, যারা অত্যন্ত গুরুত্বপূর্ণ। প্রথম জন হল **identity matrix**. এটা একটা square matrix যার প্রত্যেকটা diagonal entry-ই হল 1, আর বাকি সব entry হল 0.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

দেখতেই পাচ্ছ যে আমরা identity matrix-কে I অক্ষর দিয়ে লিখি। যদি সাইজ বোঝানোর দরকার হয়, তবে $n \times n$ সাইজের identity matrix-কে লিখি I_n.

এই matrix-দের গুরুত্ব হল এই যে, এদের দিয়ে আরেকটা matrix-কে গুণ করলে কোনো কিছুই বদলায় না। সেই দিক দিয়ে দেখলে এটা অনেকটা 1 সংখ্যাটার মত। বস্তুতঃ যদি তুমি 1×1 সাইজের identity matrix-টা (মানে I_1) লেখো, সেটা কিন্তু আসলে 1-ই হবে!

Exercise 9: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$-কে বাঁদিক থেকে I_2 আর ডানদিক থেকে I_3 দিয়ে গুণ করে দেখো তো matrix-টা সত্যিই অপরিবর্তিত থাকছে কিনা।
■

অন্য গুরুত্বপূর্ণ matrix-টা হল **zero matrix**. এটা square হতেও পারে, নাও পারে। কিন্তু সবগুলো entry-ই শূন্য। একে আমরা লিখি O. বলাই বাহুল্য যে এটা দিয়ে কোনো matrix-কে গুণ করলে ফের আরেকটা zero matrix-ই পাবে। Zero matrix মানে "সবগুলো entry-ই zero.” অনেক সময়ে **nonzero matrix** কথাটা শুনবে। এর মানে কিন্তু "সবগুলো entry-ই nonzero” নয়! এর মানে "অন্ততঃ একটা entry আছে যেটা nonzero.” যেমন $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ আর $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ দুজনেই nonzero matrix.

2.4 Block matrix operations

ধরো একটা বড় দেখে matrix দিলাম Fig 10-এর মত। লক্ষ কর এর মধ্যে একটা প্যাটার্ণ আছে। ঠিক যেন কয়েকটা একইরকম টাইল বসিয়ে বসিয়ে তৈরী। এরকম প্যাটার্ণ-ওয়ালা matrix-এর প্যাটার্ণটা সহজে বোঝানোর একটা কায়দা হল এটাকে **partition** করে লেখা। কায়দাটা এইরকম। প্রথমে Fig 11-এর মত করে কয়েকটা লাইন টানো। এই লাইনগুলোকে ভাবতে পারো টাইলগুলোর বর্ডার। এবার লক্ষ কর যে, এর মধ্যে একটা টাইল হল identity matrix, এবং বাকি তিনটে টাইল সমান। যদি এই সমান টাইলগুলোর প্রত্যেকটাকে B নাম দিই, তবে গোটা matrix-টা হয়ে যাবে

$$\begin{bmatrix} I & B \\ B & B \end{bmatrix}, \text{ যেখানে } B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

বুঝতেই পারছ এই কায়দায় প্যাটার্ণ-ওয়ালা বড় matrix-দের খুব সংক্ষেপে প্রকাশ করা যায়। খালি মনে রেখো যে, টাইলের বর্ডার হিসেবে যে লাইনগুলো টানছ, সেগুলো যেন এ মাথা থেকে ও মাথা পর্যন্ত যায়। মানে Fig 12-এর মত হলে চলবে না, কিন্তু Fig 13-এর মত হলে ঠিক আছে। এরকম টাইল দিয়ে তৈরী matrix-দের বলে **block matrix** বা **partitioned matrix.** প্রত্যেকটা টাইলকে বলে একেকটা **block.** এদের সবার সাইজ এক নাও হতে পারে।
দুটো partitioned matrix-কে অনায়াসে যোগ বিয়োগ করা যায়। যেমন

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \end{bmatrix}.$$

অবশ্য দুটোকেই একইভাবে partition করতে হবে, মানে প্রতি (i,j)-র জন্যই A_{ij} আর আর B_{ij}-র সাইজ সমান হতে হবে। এবার আমরা partitioned matrix-দের গুণ নিয়ে আলোচনা করব। মজার কথা হচ্ছে, partitioned matrix-দের প্রতিটা block-কে যদি একটা করে সংখ্যা বলে ভাবো, তবে সাধারণ matrix-দের গুণের কায়দাটা সরাসরি কাজ করে! যেমন

Fig 10

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Fig 11

$$\left[\begin{array}{cc:cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ \hdashline 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right]$$

Fig 12

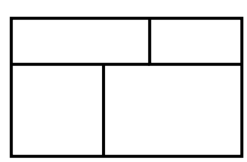
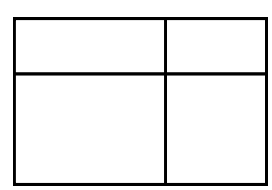


Fig 13



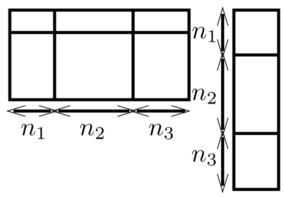


Fig 14

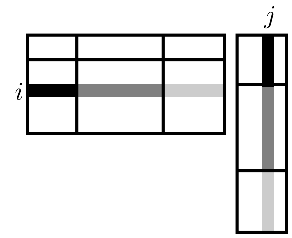


Fig 15

ধরো, আমরা জানি দুটো 2×2 matrix-কে কী করে গুণ করতে হয়--

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

ঠিক একই নিয়ম খাটবে partitioned matrix-দের বেলাতেও

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

এর জন্য অবশ্যই block-গুলোর সাইজ ঠিক মত নিতে হবে, যাতে $A_{11}B_{11}$ ইত্যাদি গুণগুলো করা যায়। সাইজগুলো ঠিক রাখার কায়দাটা খুবই সোজা--প্রথম matrix-টার column-গুলোকে যেভাবে block-এ block-এ ভাঙবে দ্বিতীয় matrix-টার row-গুলোকেও ঠিক একইভাবে ভাঙতে হবে, ব্যস্, আর কিছুই নয়। একটা উদাহরণ দিলাম Fig 14-এ। যদি প্রশ্ন জাগে যে এমনটা কী করে হচ্ছে, তবে সেটা বুঝতে পারবে এইভাবে--ধরো (i, j)-th entry-টা বার করতে চাই। তার জন্য প্রথম matrix-টার i-th row-এর সঙ্গে দ্বিতীয় matrix-টার j-th column গুণ করতে হবে। Fig 15 দেখো। লক্ষ কর যে, partition-গুলো মানানসইভাবে করা হয়েছে বলে i-th row-এর কালো অংশটা j-th column-এর কালো অংশটার সাথে গুণ হচ্ছে, একইভাবে অন্য অংশগুলোও খাপে খাপে মিলে যাচ্ছে। দুটো কালো অংশের গুণফল থেকে আসছে প্রথম partition-দুটোর গুণফলের (i, j)-th entry, দ্বিতীয়টা থেকে পরের দুটোর, ইত্যাদি। Partitioned matrix-এর গুণ নানা জায়গায় আমাদের কাজে লাগবে। তার মধ্যে বিশেষভাবে উল্লেখযোগ্য হল দুটো ক্ষেত্র। ধরো AB বার করছি।

- যদি A-কে আস্ত রাখি (মানে পুরোটাকে একখানা block বলে ভাবি), আর B-র প্রতিটা column-কে একেকটা block নিই, মানে

  $$B = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix},$$

  যেখানে $\vec{b}_j$-রা হল B-এর column-গুলো (Fig 16), তবে

  $$AB = A \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} = \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_p \end{bmatrix}.$$

  এটাকে মনে রাখতে পারো এই ভেবে যে "A-টা বাঁদিক থেকে B-এর ভিতরে ঢুকে গেল"।

Fig 16

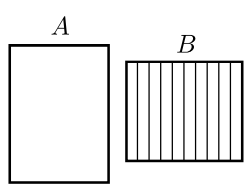
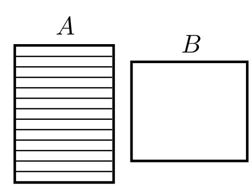


Fig 17



- এবার ধরো B-এ আস্ত রেখে A-র প্রতিটা row-কে একেকটা block নিলাম (Fig 17), মানে

$$A = \begin{bmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_m' \end{bmatrix}.$$

লক্ষ কর $\vec{a}_i$-দের transpose লিখেছি, কারণ ওরা row, মানে শোয়ানো vector, খাড়া নয়। তাহলে পাব

$$AB = \begin{bmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_m' \end{bmatrix} B = \begin{bmatrix} \vec{a}_1' B \\ \vdots \\ \vec{a}_m' B \end{bmatrix}.$$

এখানে যেন "B ডানদিক থেকে ঢুকে গেল A-র মধ্যে"।

এই দুটো বিশেষ ক্ষেত্র থেকে এই শিক্ষালাভ করা গেল যে--

- AB-র একটা column বার করতে হলে পুরো A-টা লাগে, কিন্তু B-এর খালি সেই column-টা জানলেই চলে।
- আবার AB-র কোনো row বার করতে হলে B-র পুরোটাই লাগে, কিন্তু A-র খালি সেই row-টা জানাই যথেষ্ট।

Exercise 10: দেওয়া আছে যে, $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. আরেকটা matrix আছে $B_{3\times 3}$, যার খালি দ্বিতীয় column-টাই জানা আছে (বাকি entry-গুলো জলের বোতল উল্টে মুছে গেছে!)--

$$B = \begin{bmatrix} ? & 4 & ? \\ ? & -1 & ? \\ ? & 2 & ? \end{bmatrix}.$$

এই তথ্যের ভিত্তিতে AB-র কতটা বার করতে পারবে? ■

DAY 3
নতুন জিনিস

3.1 Matrix-এর inversion

আমরা matrix-এর যোগ, বিয়োগ আর গুণের কথা শিখলাম। স্বাভাবিকভাবেই প্রশ্ন জাগে matrix-এর বেলায় ভাগও করা যায় কিনা। উত্তর হল--হ্যাঁ, কখনো কখনো যায়। ব্যাপারটা বোঝার জন্য প্রথমে সাধারণ সংখ্যার ভাগ কী করে করা হয় দেখি। যদি $2 \div 3$ বার করতে হয়, তবে তাকে আমরা লিখতে পারি $2 \times \frac{1}{3}$. এখানে $\frac{1}{3}$-কে বলে 3-এর **reciprocal**. সুতরাং কোনো সংখ্যা দিয়ে ভাগ মানে হল তার reciprocal দিয়ে গুণ করা। Matrix-এর বেলায় গুণ করা তো জানিই, এবার তবে দেখি matrix-দের বেলায় reciprocal বার করার মত কিছু করা যায় কিনা। 3-এর reciprocal হল $\frac{1}{3}$, কারণ $3 \times \frac{1}{3} = 1$. আমরা আগেই বলেছি যে, matrix-দের বেলায় 1-এর ভূমিকা পালন করে identity matrix. সুতরাং কোনো matrix-এর reciprocal-এর একটা সঙ্গত সংজ্ঞা মনে হচ্ছে এইভাবে দেওয়া যায়--

> A-র reciprocal বলব B-কে, যদি $AB = I$ হয়।

কিন্তু তাতে আবার একটা সমস্যা আছে। সাধারণ সংখ্যার ক্ষেত্রে গুণ ছিল commutative, মানে $ab = ba$. কিন্তু matrix-এর বেলায় সেটা নাও হতে পারে। তাই আলাদা করে বলে দেওয়া ভালো যে আমরা দুটো দিকই চাই, মানে $AB = I$ এবং $BA = I$. আর একটা কথা, matrix-এর বেলায় আমরা reciprocal শব্দটা ব্যবহার করি না, বলি **inverse**. তাহলে পুরো জিনিসটা দাঁড়াবে এইরকম--

DEFINITION: Inverse

Let A be a square matrix. Then a matrix B is called its **inverse** if

$$AB = I \text{ and } BA = I.$$

We write inverse of A as A^{-1}.

Example 7: ধরো $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$, $B = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$ আর $C = \begin{bmatrix} 1 & -2 \\ -3 & 5 \end{bmatrix}$. পরীক্ষা করে দেখো এদের মধ্যে কে কার inverse.

SOLUTION: এখানে $AB = I$ এবং $BA = I$. সুতরাং B হল A-র inverse (সুতরাং A আবার B-এর inverse)। কিন্তু

$$AC = \begin{bmatrix} -5 & 8 \\ -12 & 19 \end{bmatrix} \text{ আর } CA = \begin{bmatrix} -5 & -8 \\ 12 & 19 \end{bmatrix},$$

যারা কেউই I নয়। সুতরাং A এবং C পরস্পরের inverse নয়। একইভাবে দেখতে পাবে যে, B এবং C-ও পরস্পরের inverse হচ্ছে না (গুণ করেই দেখো!)। ■

Exercise 11: যদি $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ হয়, আর $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ নিই, তবে দেখো তো a, b, c, d-র value কী কী নিলে $B = A^{-1}$ হয়।

HINT:

$$AB = \begin{bmatrix} a + 2c & b + 2d \\ 3a + 4c & 3b + 4d \end{bmatrix}.$$

এটাকে I-এর সমান হতে হবে, মানে $a + 2c = 1$ হতে হবে, $2d + b = 0$ হতে হবে, ইত্যাদি। এইরকম চারটে equation পাবে। তাদেরকে solve করতে পারো কিনা দেখো। ■

Exercise 12: যদি $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ হয়, তবে কি A^{-1} বার করা যাবে? ■

Exercise 13: যদি $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ হয়, তবে কি A^{-1} বার করা যাবে? ■

এইখানে বেশ কয়েকটা কথা বলা দরকার--

1. সব matrix-এরই inverse পাওয়া যাবে এমন নয়। একটা উদাহরণ তো দেখাই যাচ্ছে, zero matrix. ওকে তুমি যা দিয়েই গুণ কর না কেন, উত্তর zero matrix-ই থাকবে, কোনোদিনই identity হবে না।

2. যদি কোনো matrix-এর একটা inverse পাও, তবে জানবে যে সেটাই একমাত্র inverse. একই matrix-এর একাধিক inverse থাকা অসম্ভব। সেই কারণে ইংরাজিতে লেখার সময়ে আমরা "*the* inverse of a matrix" লিখে থাকি। কেন একাধিক inverse থাকা অসম্ভব সেটা দেখা খুবই সহজ--ধরো A-র দুটো inverse পাওয়া গেছে B আর C. তাহলে $AC = I$ আবার $BA = I$. সুতরাং আমরা লিখতে পারি--

$$B = BI = B(AC) = (BA)C = IC = C.$$

তার মানে B আর C আসলে একই জিনিস।

3. সংজ্ঞাটায় তিনটে শর্ত আছে--matrix-টাকে square হতে হবে, $AB = I$ হতে হবে, $BA = I$ হতে হবে। এই তিনটে শর্তের মধ্যে যেকোনো দুটো থাকাই আসলে যথেষ্ট। যেমন

   - যদি $AB = I$ আর $BA = I$ হয়, তবে A-কে square হতেই হবে।
   - যদি A একটা square matrix হয় এবং এমন B পাওয়া যায় যাতে $AB = I$ হয়, তবে $BA = I$ না হয়ে যায় না,
   - একইভাবে যদি A-টা square হয় আর $BA = I$ হয়, তবে $AB = I$ হবেই।

এই তিনটের প্রমাণ কিন্তু খুব সোজা নয়। প্রমাণগুলো আমরা তৃতীয় অধ্যায়ে গিয়ে শিখব।

যেসব matrix-এর inverse থাকে, তাদের বলে **invertible**. এদের আরেকটা নাম হল **nonsingular**. যেসব square matrix-এর inverse থাকে না, তাদেরকে বলে **non-invertible** বা **singular**. একটা matrix যদি square-ই না হয়, তবে তো আর inverse থাকার প্রশ্নই ওঠে না, তাই ওদের বেলায় আমরা non-invertible বা singular কথাটা ব্যবহারই করি না।

DEFINITION: (Non)invertible/(non)singular

A square matrix having an inverse is called an **invertible** or **nonsingular** matrix. A square matrix that does not have any inverse is called a **noninvertible** or **singular** matrix.

Example 8: যদি $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ আর $B = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ হয়, তবে এদের মধ্যে একজন nonsingular আর অন্যজন singular. কোন্টা কে?

SOLUTION: A হল nonsingular, আর B হল singular. দুজনের inverse বার করার চেষ্টা করলেই দেখবে A-র বেলায় কাজটা করা যাচ্ছে, কিন্তু B-র বেলায় যাচ্ছে না। ■

3.1.1 2×2 সাইজের matrix-এর inverse

কোনো square matrix দেওয়া থাকলে চট্ করে সেটা nonsingular কিনা সেটা বলতে পারা, এবং nonsingular হলে তার inverse বার করা সহজ নাও হতে পারে। কিন্তু 2×2 সাইজের matrix-দের বেলায় একটা সহজ কায়দা আছে যেটা জেনে রাখা ভালো--

THEOREM

A matrix $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is nonsingular if and only $ad-bc\neq 0$, and in this case

$$A^{-1}=\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

এটা কেন হয়, সেটা বোঝা খুবই সহজ। যদি $ad-bc\neq 0$ হয়, তবে এই A^{-1}-এর সঙ্গে A-কে গুণ করে দিলেই দেখবে I পেয়ে যাচ্ছ (করেই দেখো!), সুতরাং A^{-1} বলে যেটা দিয়েছি, সেটা সত্যিই A^{-1}. আর nonsingular হবার জন্য কেন $ad-bc\neq 0$ হতে হবে, সেটা এইভাবে দেখা যায়। যদি A-টা nonsingular হয়, তবে ধরো

$$A^{-1}=\begin{bmatrix} p & q \\ r & s \end{bmatrix}.$$

সুতরাং

$$A^{-1}A=\begin{bmatrix} p & q \\ r & s \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

এ থেকে তুমি চারটে equation পাবে (যেহেতু matrix-গুলোয় চারটে করে entry আছে)। যেমন $(1,1)$-th entry-র জন্য পাবে

$$pa+qc=1.$$

এই চারটে equation থেকে যদি p,q,r,s বার করতে যাও, তবেই দেখবে $ad-bc\neq 0$ শর্তটা এসে যাচ্ছে (যদিও ধাপগুলো করতে গিয়ে মাথা একটু গুলিয়ে যেতে পারে)। তোমরা হায়ার সেকণ্ডারী থেকেই জানো যে $ad-bc$-কে এখানে determinant বলে। এই বইতে আমরা পরে determinant নিয়ে বিস্তারিত আলোচনা করব।

3.2 বিভিন্ন matrix operation-এর ধর্ম

Matrix-দের নিয়ে কয়েকরকম কাজ করা আমরা শিখেছি, যোগ, বিয়োগ, গুণ, transpose আর inverse বার করা। এদেরকে বলে একেকটা **matrix operation**. এই operation-গুলোর মধ্যে বিভিন্ন ধর্ম জানা থাকলে কাজ করা সহজ হয়। এদের মধ্যে কয়েকটা আমাদের অতি পরিচিত--

- $A+B=B+A$
- $(A+B)+C=A+(B+C)$
- $(AB)C=A(BC)$
- $A(B+C)=AB+AC$
- $(B+C)A=BA+CA$

এইগুলো ঠিক সাধারণ সংখ্যার যোগ বিয়োগের মতই। খালি দুটো জিনিস আছে যেখানে সামান্য হোঁচট খেতে হয়--

- AB আর BA সমান নাও হতে পারে।
- $AB=O$ হলেই কিন্তু বলা যায় না যে, হয় $A=O$ নয়তো $B=O$ হবেই। একইভাবে $A^2=O$ মানেই যে $A=O$ হবে, এমন কোনোই কথা নেই। এটা একটা গল্প দিয়ে ভাবা যায়। আমরা একটু আগে রামায়ণের অংকটা করেছিলাম মনে আছে? সেখানে একটা matrix ছিল "কে কার বাবা?" ধরো এটার নাম দিলাম A. তাহলে বুঝতেই

পারছ A^2 হবে "বাবার বাবা" মানে "'ঠাকুর্দা"-র matrix. আমাদের উদাহরণে চরিত্র ছিল দশরথ, কৌশল্যা, রাম, সীতা, জনক। এখানে কেউ কারো ঠাকুর্দা নয়, সুতরাং $A^2 = O$ হতে বাধ্য (সন্দেহ থাকলে গুণটা করে মিলিয়েই নাও!)। একটা ঘরের মধ্যে কয়েকজন বাবা থাকলেই যে কিছু ঠাকুর্দাও থাকতে হবে এমন তো কথা নেই, তেমনি $A \neq O$ হলেই A^2-কেও nonzero হতে হবে এমন কোনো কথা নেই! এই উদাহরণে হাতেকলমে A^2 বার করলে দেখবে যে সব সময়েই nonzero entry-দের সঙ্গে 0 গুণ হয়ে যাচ্ছে। যদি দুটো matrix-এর প্রতিটা entry-ই nonzero হয়, তাও কিন্তু matrix দুটো গুণ করলে product-টা একটা zero matrix হতেই পারে। যেমন নীচের গুণটায় দুটো matrix-এই কোথাও কোনো 0 নেই, তাও গুণফলটা O হয়েছে--

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

এর পরে আসি transpose আর conjugate transpose-এর কথায়--

- $(A')' = A$ এবং $(A^*)^* = A$.
- $(A+B)' = A' + B'$ এবং $(A+B)^* = A^* + B^*$.
- $(AB)' = B'A'$ এবং $(AB)^* = B^*A^*$.

লক্ষ কর যে শেষের ক্ষেত্রে (conjugate) transpose নেওয়ার ফলে B-টা কেমন আগে চলে এল। এই ক্রম উল্টে যাওয়ার ব্যাপারটা দুয়ের বেশী matrix-এর বেলাতেও হয়, যেমন $(ABCD)' = D'C'B'A'$ এবং $(ABCD)^* = D^*C^*B^*A^*$ হবে।

সব শেষে আসি inverse-এর কথায়। এখানে কিছু জিনিস বেশ সহজ বুদ্ধিতেই বোঝা যায়, আর কোথাও কোথাও হোঁচট খেতে হয়। ভালো করে বুঝে বুঝে পড়।

- যদি A হয় nonsingular, তবে $\left(A^{-1}\right)^{-1} = A$ হবে। ঠিক যেমন $\frac{1}{\frac{1}{2}} = 2$ হয়।
- যদি A, B দুজনেই একই সাইজের দুটো square matrix হয়, তবে AB-টা nonsingular হলেই A আর B দুজনেই আলাদা করে nonsingular হবে, এবং conversely, যদি A, B দুজনেই nonsingular হয়, তবে AB-ও nonsingular হতে বাধ্য। এবং দুইক্ষেত্রেই $(AB)^{-1} = B^{-1}A^{-1}$ হবে। এটাকে মনে রাখার একটা ভালো কায়দা আছে। মনে করো A হল "মোজা পরা" আর B হল "জুতো পরা" তাহলে AB-কে ভাবা যেতে পারে মোজা পরার পর জুতো পরা। এবার যদি তোমাকে মোজা-জুতো খুলে ফেলতে হয়, তবে তুমি নিশ্চয়ই আগে জুতো খুলবে (মানে B^{-1}) এবং তারপরে মোজা খুলবে (মানে A^{-1}). পরার সময়ে যেই ক্রমে কাজগুলো করা হয়েছিল, খোলার সময়ে সেই ক্রমটা উল্টে যাবে।

  এটার একটা দিক প্রমাণ করা খুবই সোজা--A, B দুজনেই nonsingular হলে AB-ও nonsingular হবে। সরাসরি গুণ করেই দেখো--

  $$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

  উল্টো দিকটা কিন্তু অত সহজ নয়--মানে AB যদি nonsingular হয়, তবে A, B দুজনেই কেন আলাদা করে nonsingular হবে সেটা বোঝার জন্য আরও যন্ত্রপাতি লাগবে। সেসব আমরা তৃতীয় অধ্যায়ে গিয়ে দেখব।

- যদি A একটা nonsingular matrix হয় আর $a \neq 0$ কোনো সংখ্যা হয়, তবে aA-ও হবে nonsingular, এবং

  $$(aA)^{-1} = \tfrac{1}{a}A^{-1}$$

  হবে।

- যদি A আর B দুজনেই সমান সাইজের nonsingular matrix হয়, তবে কি $A+B$-ও nonsingular হবে? উঁহু! ধরো $A = I$ আর $B = -I$, তবে $A + B = O$, যেটা কোনোভাবেই nonsingular নয়! আর যদি কখনও $A+B$ দেখো যে nonsingular হয়েও যায়, তাও যেন খবরদার ভেবে বোসো না যে $(A+B)^{-1}$ হবে $A^{-1} + B^{-1}$.

- অবশ্য transpose-এ সঙ্গে inverse-এর গলায় গলায় ভাব। কীরকম শোনো। ধরো একটা matrix আছে A. যদি A-টা nonsingular হয় তবে A'-ও nonsingular হবেই। আর সেক্ষেত্রে $(A')^{-1} = \left(A^{-1}\right)'$ হতে বাধ্য। মানে, inverse-এর transpose আর transpose-এর inverse একই কথা। এর প্রমাণটা খুবই সহজ। সেটাই চেয়েছে নীচের অংকটায়।

Example 9: If A is an invertible matrix, prove that A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$, where A^t denotes the transpose of A.[5] **(2002)**

SOLUTION:

> Since A is invertible, A^{-1} exists and
>
> $$AA^{-1} = I \text{ and } A^{-1}A = I.$$
>
> Taking transpose of both sides,
>
> $$(AA^{-1})^t = I^t \text{ and } (A^{-1}A)^t = I^t,$$
>
> or
>
> $$(A^{-1})^t A^t = I \text{ and } A^t (A^{-1})^t = I, \qquad (*)$$

এখানে কী কী তথ্য ব্যবহার করলাম লিখে নিই--

> Since I is symmetric, and for any two matrices B, C for which BC is defined, we have $(BC)^t = C^t B^t$.
> (*) shows that A^t is invertible with
>
> $$(A^t)^{-1} = (A^{-1})^t,$$
>
> as required.

■

Transpose-এর জন্য যা বলা গেল conjugate transpose-এর বেলাতেও কি তাই বলা যাবে? নীচের অংকটা বলছে--হ্যাঁ, যাবে।

Exercise 14: ধরো A একটা nonsingular matrix. তবে দেখাও যে $\left(A^{-1}\right)^* = \left(A^*\right)^{-1}$. ■

একটা matrix তৈরী হয় অনেকগুলো সংখ্যা দিয়ে, কিন্তু পুরোটা মিলে বেশ একখানা সংখ্যার মত আচরণ করে। নীচের অংকটা সেই সাদৃশ্যটাকেই তুলে ধরেছে।

Example 10: Find the 3×3 matrices A and B when $A + B = 2B^t$ and $2A + 3B = 6I_3$, with

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

[3] **(2004)**

SOLUTION:

Given

$$A + B = 2B^t \qquad (\mathbf{1})$$

and

$$2A + 3B = 6I. \qquad (\mathbf{2})$$

এখানে মনে করো যেন A, B দুটো সংখ্যা, যেগুলোকে equation দুটো solve করে বার করতে হবে। প্রথমে A-কে B দিয়ে লিখে নিই--

From $(\mathbf{2})$

$$A = 3I - \tfrac{3}{2}B.$$

Putting this in $(\mathbf{1})$ we have $3I - \frac{3}{2}B + B = 2B^t$, or

$$3I = \tfrac{1}{2}B + 2B^t. \qquad (\mathbf{3})$$

এর মধ্যে B ছাড়া কিছু নেই বটে, কিন্তু B^t এসে গিয়ে সমস্যা হয়েছে। তাই সামান্য কৌশল করব--

Transposing,

$$3I = \tfrac{1}{2}B^t + 2B \quad (\because I^t = I). \qquad (\mathbf{4})$$

Now solving $(\mathbf{3})$ and $(\mathbf{4})$ we get

$$B = \tfrac{6}{5}I,$$

and

$$A = 3I - \tfrac{3}{2}B = \tfrac{6}{5}I.$$

■

এবার একটা অংক করব যেটা বেশ সহজ, কিন্তু একটা কৌশল না জানলে দেখতে খুব কঠিন লাগে। কৌশলটা অনেক জায়গাতেই কাজে লাগে। তাই একটু আলোচনা করে নিই। ধরো একটা matrix দিলাম--

$$A = \begin{bmatrix} 1 & 2 & 9 \\ 3 & 4 & -4 \end{bmatrix},$$

এবং বললাম এর প্রতিটা row-এর entry-গুলোর যোগফল বার করতে। এখানে প্রথম row-এর তিনটে সংখ্যার যোগফল হল $1+2+9 = 12$ আর দ্বিতীয় row-এর বেলায় $3+4+(-4) = 3$. এদের দিয়ে একটা column vector বানানো যাক--

$$\vec{b} = \begin{bmatrix} 12 \\ 3 \end{bmatrix}.$$

A থেকে $\vec{b}$ বার করার কায়দাটাকে সুন্দর করে matrix multiplication দিয়ে লেখা যায় এইভাবে--

$$\begin{bmatrix} 12 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 9 \\ 3 & 4 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

ভালো করে বুঝে নাও ব্যাপারটা কী হচ্ছে। সবার শেষে ওই সব 1 দিয়ে তৈরী vector-টা রয়েছে ওটা খুব কাজের জিনিস, এবং আমাদের কৌশলটার মূল উপাদান। ওটাকে যদি $\vec{\mathbb{1}}$ লিখি তবে, "row-এর entry-গুলোর sum বার করা" কাজটাকে লিখতে পারি

$$\vec{b} = A\vec{\mathbb{1}}.$$

একইভাবে যদি column-এর entry-গুলোর sum বার করতে চাইতাম, তবে লিখতাম

$$\left[\begin{array}{cc} 1 & 1 \end{array}\right] \left[\begin{array}{ccc} 1 & 2 & 9 \\ 3 & 4 & -4 \end{array}\right] = \vec{\mathbb{1}}'A.$$

এখানে বাঁদিক থেকে গুণ করেছি একটা শোয়ানো $\vec{\mathbb{1}}$ দিয়ে। এখানে $\vec{\mathbb{1}}$-এর দৈর্ঘ্য হল 2. দেখতেই পাচ্ছ $\vec{\mathbb{1}}$ চিহ্নটা যে কোনো দৈর্ঘ্যের vector-এর জন্যই ব্যবহার করছি, খালি প্রতিটা entry-ই যেন 1 হয়।
এইবার নীচের অংকটায় হাত দেওয়া যাক।

Example 11: A is a nonsingular matrix such that the sum of the elements in each row is K.

Prove that the sum of the elements in each row of A^{-1} is K^{-1}.[3] **(2013.1bi)**
SOLUTION:

> Let A be $n \times n$ and let $\vec{\mathbb{1}} = \left[\begin{array}{c} 1 \\ \vdots \\ 1 \end{array}\right]_{n\times 1}$.
>
> Then we have
> $$A\vec{\mathbb{1}} = K\vec{\mathbb{1}}. \qquad (*)$$

এটা ভালো করে বুঝে নাও। দরকার হলে একটা উদাহরণ নাও, ধরো $A = \left[\begin{array}{cc} 1 & 4 \\ 3 & 2 \end{array}\right]$. এর row-এর সংখ্যাগুলো যোগ করলে দুটো row-এর বেলাতেই যোগফল হয় 5. তাই আমরা বলতে পারি

$$\left[\begin{array}{cc} 1 & 4 \\ 3 & 2 \end{array}\right] \left[\begin{array}{c} 1 \\ 1 \end{array}\right] = \left[\begin{array}{c} 5 \\ 5 \end{array}\right] = 5\left[\begin{array}{c} 1 \\ 1 \end{array}\right],$$

মানে $A\vec{\mathbb{1}} = 5\vec{\mathbb{1}}$.
একই কায়দায় লিখে নিই, কী দেখাতে হবে--

> To show $A^{-1}\vec{\mathbb{1}} = \frac{1}{K}\vec{\mathbb{1}}$.

প্রমাণটা খুবই সংক্ষিপ্ত--

> $\because A$ is nonsingular, $\therefore A^{-1}$ exists.
>
> Multiplying (*) by A^{-1} from left, we have
> $$A^{-1}A\vec{\mathbb{1}} = KA^{-1}\vec{\mathbb{1}},$$
> or
> $$\vec{\mathbb{1}} = KA^{-1}\vec{\mathbb{1}}.$$
> Now K cannot be 0, because then the RHS is zero, but the LHS is nonzero.

So we can divide both sides by K to get

$$A^{-1}\vec{\mathbb{1}} = \frac{1}{K}\vec{\mathbb{1}}.$$

So each row sum of A^{-1} is $\frac{1}{K}$, as required.

■

নীচের অংকটা একই কায়দায় করতে পারবে।

Exercise 15: A is a nonsingular matrix such that the sum of the elements in each column is K. Prove that the sum of the elements in each column of A^{-1} is $\frac{1}{K}$.
HINT: এখানে $\vec{\mathbb{1}}'A = K\vec{\mathbb{1}}'$. এবার এগোও। ■

স্কুলে যেভাবে এমনি সংখ্যাদের অংক করতাম, matrix নিয়েও প্রায় সেভাবেই করা যায়, খালি গুণটা commutative না হওয়ায় অনেক সময়ে ফাঁপরে পড়ে যেতে হয়। তখন বাঁদিকের জিনিসকে ডানদিকে আনতেই অনেকগুলো ধাপ খরচ করতে হয়। নীচের অংকটা তার একটা নমুনা।

Example 12: If A and B are two $n \times n$ matrices and A has an inverse, then show that

$$(A+B)A^{-1}(A-B) = (A-B)A^{-1}(A+B).$$

[2] **(2001)**
SOLUTION: আহা, যদি matrix-এর গুণ commutative হত, তবে এর মধ্যে দেখানোর কিছুই থাকত না! এখানে LHS আর RHS দুটোই বেশ বিচ্ছিরি দেখতে। তাই দুটোকেই প্রথমে যতটা সম্ভব সহজ করে নিই--

The LHS is

$$\begin{aligned}(A+B)A^{-1}(A-B) &= AA^{-1}A - AA^{-1}B + BA^{-1}A - BA^{-1}B \\ &= A - B + B - BA^{-1}B \\ &= A - BA^{-1}B.\end{aligned}$$

Similarly, the RHS is

$$\begin{aligned}(A-B)A^{-1}(A+B) &= AA^{-1}A + AA^{-1}B - BA^{-1}A - BA^{-1}B \\ &= A + B - B - BA^{-1}B \\ &= A - BA^{-1}B = \text{LHS},\end{aligned}$$

as required.

■

আরেকটা অংক দেখি। এখানেও commutative না হওয়া নিয়ে সমস্যা।

Example 13: If A is a nonsingular matrix show that a matrix B commutes with A^{-1} if and only if B commutes with A. **(2007)**

SOLUTION: এখানে একটা নতুন ভাষা ব্যবহার করেছে "commutes with". যদি A, B দুটো matrix হয় এবং বলা থাকে যে, B commutes with A, তার মানে হল $AB = BA$. এবার তাহলে দেখি এই অংকটাতে কী চেয়েছে।

To show

$$AB = BA \iff A^{-1}B = BA^{-1}.$$

$(\implies)$:

$$\begin{aligned} AB = BA &\implies A^{-1}AB = A^{-1}BA \quad \left[\text{multiplying by } A^{-1} \text{ from left}\right] \\ &\implies B = A^{-1}BA \\ &\implies BA^{-1} = A^{-1}BAA^{-1} \quad \left[\text{multiplying by } A^{-1} \text{ from right}\right] \\ &\implies BA^{-1} = A^{-1}B, \end{aligned}$$

as required.

এবার অন্য দিকটা। যুক্তিটা এখানেও একই, খালি A-র জায়গায় A^{-1} বসালেই হবে।

$(\impliedby)$:

$$\begin{aligned} BA^{-1} = A^{-1}B &\implies BA^{-1}A = A^{-1}BA \quad \left[\text{multiplying by } A \text{ from right}\right] \\ &\implies B = A^{-1}BA \\ &\implies AB = AA^{-1}BA \quad \left[\text{multiplying by } A \text{ from left}\right] \\ &\implies AB = BA, \end{aligned}$$

as required.

■

DAY 4
নানা রকমের matrix (part 1)

Matrix নানারকমের হয়। এদের মধ্যে কয়েকধরণের matrix-এর জনপ্রিয়তা খুবই বেশী। এই জনপ্রিয়তার কারণ মূলতঃ একটাই। যোগ, বিয়োগ, গুণ ইত্যাদির সঙ্গে যেসব ধরণের matrix-এর সম্পর্ক ভালো তারাই বেশী জনপ্রিয় হয়। "সম্পর্ক ভালো" বলতে কী বোঝাচ্ছি একটা উদাহরণ দিই। মনে কর diagonal matrix-দের কথা দিয়ে শুরু করি।

4.1 Diagonal matrix

এরা সবাই square হয়, এবং diagonal-এর বাইরে সব entry-ই হয় শূন্য। Diagonal entry-গুলোও কেউ কেউ শূন্য হলে আপত্তি নেই। কয়েকটা উদাহরণ--

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

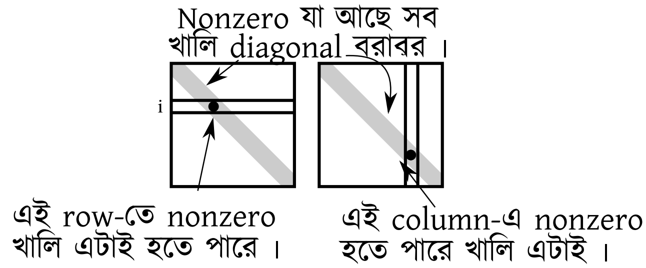


Fig 18

এরকম দুটো matrix নাও, একই সাইজের। দুটোকে যোগ বা বিয়োগ করে দেখো উত্তরটা ফের diagonal matrix-ই হবে। এবার গুণ করে দাও, তাহলেও diagonal matrix-ই থাকবে। শুধু তাই নয় diagonal entry-গুলো গুণ হয়ে যাবে। নীচের গুণটা হাতেকলমে করে দেখলেই কথাটার সত্যতা মালুম হবে--

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 10 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 30 \end{bmatrix} = \begin{bmatrix} 1\times 10 & 0 & 0 \\ 0 & 2\times 20 & 0 \\ 0 & 0 & -1\times 30 \end{bmatrix} = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 40 & 0 \\ 0 & 0 & -30 \end{bmatrix}.$$

যারা সব কিছুকে অংকের ভাষায় লিখে প্রমাণ করতে ভালোবাসো, তারা এইভাবেও ব্যাপারটাকে দেখতে পারো-- দুটো diagonal matrix নিচ্ছি A আর B. মানে $i \neq j$ হলে $a_{ij} = b_{ij} = 0$ হবে। এদের গুণ করে পাচ্ছি $C = AB$, তার মানে

$$c_{ij} = \sum_k a_{ik} b_{kj}.$$

এই ফর্মুলাটা কী করে লিখতে হয়, সেটা আগেই বলেছি। এর মধ্যে যে a_{ik} আর b_{kj} আছে, এদের nonzero হবার আশা একমাত্র $i = k$ এবং $k = j$ হলেই। তার মানে c_{ij}-কে যদি nonzero হতে হয়, তবে $i = k = j$ দরকার। Fig 18-এর দিকে চোখ রাখো। সুতরাং $i \neq j$ হলে $c_{ij} = 0$ হতে বাধ্য, মানে C হবে একটা diagonal matrix. যদি $i = j$ হয়, তবে sum-টায় খালি একটা term-ই থাকবে (যখন $k = i = j$), সেটা হল

$$c_{ii} = a_{ii} b_{ii},$$

মানে diagonal entry-রা গুণ হয়ে যাচ্ছে।

Diagonal matrix-দের আরেকটা বৈশিষ্ট্য আছে, সেটার কথা একটু না বললেই নয়। তুমি যা খুশি একটা matrix নিয়ে সেটাকে বাঁদিক থেকে একটা diagonal matrix দিয়ে গুণ কর, যেমন ধরো

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 12 & 15 & 18 \end{bmatrix}.$$

এর মধ্যে একটা প্যাটার্ণ দেখতে পাচ্ছ? এখানে diagonal matrix-টার প্রথম diagonal entry-টা ছিল 2. লক্ষ কর যে অন্য matrix-টার পুরো প্রথম row-টা এই 2 দিয়ে গুণ হয়ে গেছে! একইভাবে দ্বিতীয় row-টা পুরো গুণ হয়ে গেছে 3 দিয়ে, কারণ diagonal matrix-টার দ্বিতীয় diagonal entry-টা ছিল 3.

Exercise 16: এমনটা যে সবসময়েই হবে সেটা অনায়াসে প্রমাণ করা যায় সেই কায়দাটা দিয়ে--AB-র (i, j)-th entry হবে $\sum_k a_{ik} b_{kj}$. ধরো A হল diagonal, তার মানে $i \neq j$ হলে $a_{ij} = 0$. এ থেকে তোমাকে দেখাতে হবে যে AB-র (i, j)-th entry হবে $a_{ii} b_{ij}$. কী করে করবে? ■

একইরকম একটা ব্যাপার হয় column-দের ক্ষেত্রেও, খালি তখন diagonal matrix-টাকে ডানদিকে রাখতে হবে, যেমন

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 6 & -3 \\ 8 & 15 & -6 \end{bmatrix}.$$

Exercise 17: এটাকেও আগের অংকটার মত প্রমাণ করো দেখি। ■

এই ব্যাপারটা পরেও বার বার দেখবে--কোনো matrix-এর row-গুলোর উপর প্রভাব বিস্তার করতে হলে বাঁদিক থেকে গুণ করলে কাজ হয়, আর column-দের উপর প্রভাব বিস্তার করতে চাইলে সুবিধা হয় ডানদিক থেকে গুণ করলে।
এক্ষুণি যা শিখলাম তার থেকে আরেকটা জিনিস করা যায়--একটা diagonal matrix দিলে একবার তাকিয়েই চোখের নিমেষে বলে দেওয়া যায় সেটা nonsingular কিনা, এবং যদি nonsingular হয়, তবে তার inverse-ও একবারে লিখে দেওয়া যায়! একটা উদাহরণ দেখি।

Example 14: এই diagonal matrix-টা কি nonsingular?

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

যদি হয়, তবে inverse বার করো।
SOLUTION: তার মানে আমরা খুঁজছি এমন কোনো B পাওয়া যায় কি না যাতে $AB = I$ হয়। যেহেতু A একটা diagonal matrix, এবং গুণের সময়ে A-টা বাঁদিকে আছে, তার মানে AB-র প্রথম row-টা হবে B-এর প্রথম row-এর 2-গুণ (যেহেতু A-র প্রথম diagonal entry হল 2). আমরা চাইছি $AB = I$, তার মানে AB-র প্রথম row-টা আমরা $(1, 0, 0)$ চাই। এটাকে হতে হবে B-এর প্রথম row-এর 2-গুণ, মানে B-এর প্রথম row-টা পেয়ে গেলাম--$\left(\frac{1}{2}, 0, 0\right)$. একই যুক্তিতে B-এর দ্বিতীয় row-এর -1-গুণ হবে $(0, 1, 0)$, এবং তৃতীয় row-এর 10-গুণ হবে $(0, 0, 1)$. সুতরাং দেখাই যাচ্ছে যে A একটা nonsingular matrix, যার inverse হল

$$B = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{10} \end{bmatrix}.$$

লক্ষ কর যে, B-টাও আবার একটা diagonal matrix-ই হয়েছে, এবং ওর diagonal entry-গুলো হল A-র diagonal entry-গুলোর reciprocal. ■

এবার আরেকটা উদাহরণ দেখি।

Example 15: এই diagonal matrix-টা কি nonsingular?

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

যদি হয়, তবে inverse বার করো।
SOLUTION: আগের অংকের মতই এমন B চাই যাতে $AB = I$ হয়। এখানেও A একটা diagonal matrix. সুতরাং আগের মতই এখানেও AB-র প্রথম row হবে B-এর প্রথম row-এর 0-গুণ (কারণ A-র প্রথম diagonal entry হচ্ছে 0). তার মানে B-এর প্রথম row যাই হোক না কেন AB-র প্রথম row ওই 0-টার ঠ্যালায় $(0, 0, 0)$ হয়ে যাবে। ফলে কোনোভাবেই $AB = I$ হতে পারে না। তার মানে একটা diagonal entry শূন্য থাকার ফলে A-র আর nonsingular হয়ে ওঠা হল না! ■

সব মিলিয়ে কী শিখলাম সেটা একটা theorem আকারে লিখে রাখি--

> THEOREM
>
> A diagonal matrix is nonsingular if and only if all the diagonal entries are nonzero. In this case the inverse is a diagonal matrix whose diagonal entries are the reciprocals of the corresponding entries of the original matrix.

সুতরাং দেখতেই পাচ্ছ যে, diagonal matrix নিয়ে কাজ করার কত সুবিধা। যত রকমের যাই কর না কেন matrix-টা diagonal-ই থাকে। সেই জন্যেই বলেছিলাম যে diagonal matrix-দের সঙ্গে বিভিন্ন matrix operation-এর ভারি সদ্ভাব!

4.2 Triangular matrix

Diagonal matrix-দের মত অতটা ভালো না হলেও **triangular matrix**-রাও বেশ ভালো জিনিস। এরা হল সেই সব square matrix যাদের diagonal-এর কোনো এক দিকের সব entry হল শূন্য, যেমন--

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix}.$$

প্রথমটার ক্ষেত্রে diagonal-এর নীচের entry-গুলো সবাই শূন্য (মানে nonzero যা আছে, সবাই আছে diagonal-এর উপরে)। তাই একে বলব **upper triangular matrix**. লক্ষ কর যে diagonal-এর উপরের সব entry-কেই যে nonzero হতে হবে, এমন কোনো কথা কিন্তু নেই। একই রকম যুক্তিতে দ্বিতীয় matrix-টাকে বলব **lower triangular matrix**.

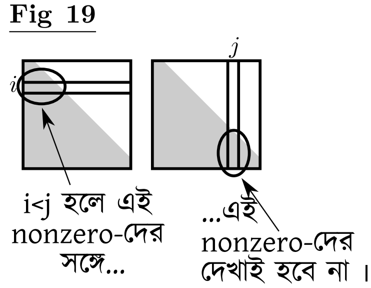
আশা করি সন্দেহ নেই যে, একটা upper triangular matrix-কে transpose করলে একটা lower triangular matrix পাওয়া যায়। তেমনি একটা lower-এর transpose নিলে upper হয়। দুটো upper triangular matrix-কে যোগ বা বিয়োগ করলে উত্তরটাও upper triangular হতে বাধ্য। একইভাবে, দুটো lower-কে যোগ-বিয়োগ করলে lower হবেই। যেটা খুব স্পষ্ট নাও হতে পারে, সেটা হল দুটো upper triangular matrix-এর গুণফলও upper triangular হয়। শুধু তাই নয়, diagonal entry-গুলোও গুণ হয়ে যাবে। একই রকম কথা অবশ্যই খাটে lower triangular-দের ক্ষেত্রেও। একটা উদাহরণ দেখি--

$$\begin{bmatrix} 4 & 0 & 0 \\ 1 & 3 & 0 \\ 6 & 2 & 10 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 8 & 5 & 0 \\ 2 & 6 & 1 \end{bmatrix} = \begin{bmatrix} 16 & 0 & 0 \\ 28 & 15 & 0 \\ 60 & 70 & 10 \end{bmatrix}.$$

লক্ষ কর diagonal entry-গুলো কেমন গুণ হয়ে গেছে, $4 \times 4 = 16,\ 3 \times 5 = 15$, ইত্যাদি। এমনটা যে হবে সেটা অংক কষে প্রমাণ করা খুবই সহজ, নীচের অংকে সেটাই করতে বলেছে।

Exercise 18: যদি A, B দুজনেই lower triangular হয়, তবে $i < j$ হলেই $a_{ij} = b_{ij} = 0$. ধরো $C = AB$. তবে দেখাও যে C-ও lower triangular হতে বাধ্য। Fig 19 দেখলে সুবিধা হতে পারে। ■

Fig 19



```
   S E N D
 + M O R E
 ---------
 M O N E Y
```

Fig 20

একটা triangular matrix কখন nonsingular হবে, সেটাও চোখে দেখেই ধাঁ করে বলে দেওয়া যায়। তার জন্য কাজে লাগে এই theorem-টা--

> **THEOREM**
>
> An upper (lower) triangular matrix A is nonsingular if and only if all its diagonal entries are nonzero. In this case the inverse is also an upper (lower) triangular matrix. Also the diagonal entries of A^{-1} are the reciprocals of the diagonal entries of A.

এই theorem-টা প্রমাণ করা কঠিন কিছু নয়। নীচের অংকটা করলেই সেটা বুঝতে পারবে। অংকটা থেকে আরেকটা জিনিসও দেখবে যে, triangular matrix-রা nonsingular হলে তাদের inverse-টা মোটামুটি সহজেই বার করে ফেলা যায়।

Exercise 19: এই triangular matrix-টার inverse থাকলে বার করো--

$$\begin{bmatrix} 2 & 0 & 0 \\ 5 & -1 & 0 \\ 6 & 0 & 10 \end{bmatrix}.$$

HINT:

ছোটোবেলায় একটা ধাঁধা দেখেছিলাম, সেখানে একটা যোগ অংক দেওয়া ছিল Fig 20-এর মত। প্রত্যেকটা অক্ষরের জায়গায় একটা করে সংখ্যা বসাতে হবে, যাতে যোগটা মিলে যায়। বিভিন্ন অক্ষরের জায়গায় বিভিন্ন সংখ্যা বসিয়ে বসিয়ে টুকটুক করে দেখতে হত। প্যাঁচটা ছিল এইটা বোঝা যে, কোথা থেকে শুরু করতে হবে। আমাদের অংকটাও অনেকটা সেইরকম। এখানে আমরা একটা B খুঁজছি যাতে $AB = I$ হয়। B-এর মধ্যে $3 \times 3 = 9$-টা সংখ্যা আছে। তাদের b_{ij} নাম দিয়ে নিলে আমরা চাইছি--

$$\begin{bmatrix} 2 & 0 & 0 \\ 5 & -1 & 0 \\ 6 & 0 & 10 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

প্রথমে গুণফলটার $(1, 1)$-th entry-টা বার করি--

$$2 \times b_{11} + 0 \times b_{21} + 0 \times b_{31} = 1.$$

এত শূন্যের ছড়াছড়ি হওয়ায় এ থেকে চট্ করে b_{11} ছাড়া সব কিছুই কেটে গিয়ে পড়ে থাকবে $2b_{11} = 1$, মানে $b_{11} = \frac{1}{2}$. সুতরাং আমরা পেলাম--

$$\begin{bmatrix} 2 & 0 & 0 \\ 5 & -1 & 0 \\ 6 & 0 & 10 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

এবার গুণফলটার $(1,2)$-th entry-র দিকে তাকালে একইভাবে পাবে $b_{12}=0$. এইভাবে এগোতে থাক, দেখবে শেষ অব্ধি আসছে

$$B=\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ \frac{5}{2} & -1 & 0 \\ -\frac{3}{10} & 0 & \frac{1}{10} \end{bmatrix}.$$

■

DAY 5
নানা রকমের matrix (part 2)

5.1 Symmetric এবং Hermitian matrix

যে সব matrix-এর transpose নিলেও একই থাকে তাদের বলে **symmetric matrix**, মানে $A'=A$. এই রকম matrix-এর leading diagonal বরাবর একটা আয়না বসালে matrix-টা যে দুটো ভাগে ভেঙে যায়, তারা আয়নায় পরস্পরের প্রতিফলন। Fig 21 দেখো। যে সব matrix-এর conjugate transpose নিলেও একই থাকে, তাদের বলে **Hermitian matrix**, অর্থাৎ $A^*=A$. যদি $A'=-A$ হয়, তবে তাকে বলে **skew symmetric**. একইভাবে $A^*=-A$ হলে বলা হয় **skew Hermitian**. বুঝতেই পারছ যে, symmetric বা Hermitian বা skew symmetric বা skew Hermitian matrix-রা square হতে বাধ্য।

Example 16: নীচের matrix-গুলোর মধ্যে কোনটা symmetric, Hermitian, skew symmetric, skew Hermitian বল।

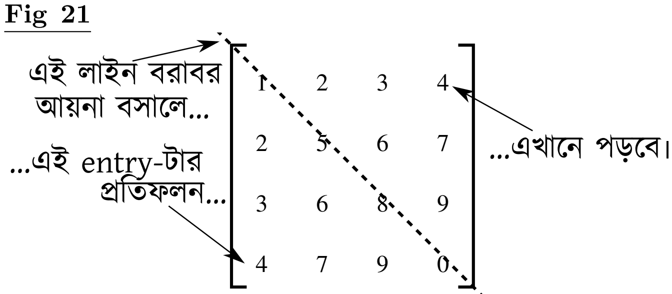
$$A=\begin{bmatrix} 1 & 2 \\ 2 & 10 \end{bmatrix},\quad B=\begin{bmatrix} 0 & 2+i \\ -2+i & 0 \end{bmatrix},\quad C=\begin{bmatrix} 1 & 2 \\ -2 & 10 \end{bmatrix}.$$

SOLUTION: এখানে A হল real, symmetric, অতএব Hermitian-ও বটে। B হল skew-Hermitian. C কোনোটাই নয়। সাবধান, C-কে যেন skew-symmetric ভেবে বোসো না! কারণ $C'=\begin{bmatrix} 1 & -2 \\ 2 & 10 \end{bmatrix}\neq -C$. ■

Symmetric matrix-দের বেলায় $a_{ij}=a_{ji}$ হবে। আর Hermitian-দের বেলায় $a_{ij}=\overline{a_{ji}}$.
একইভাবে skew symmetric-দের ক্ষেত্রে হবে $a_{ij}=-a_{ji}$ আর skew Hermitian-দের জন্য $a_{ij}=-\overline{a_{ji}}$. এই কয়টা জিনিস মাথায় রেখে নীচের অংকদুটো কর তো!

Exercise 20: যদি A একটা skew symmetric matrix হয়, তবে a_{ii} কী হবে? ■

Fig 21



Exercise 21: যদি A একটা skew Hermitian matrix হয়, তবে a_{ii} কী হবে? ■

যদি দুটো symmetric matrix নিয়ে যোগ বা বিয়োগ কর, তবে উত্তরটাও symmetric হবে। একই রকম কথা খাটে Hermitian matrix-দের ক্ষেত্রেও। এদের skew সংস্করণদের বেলাতেও একই ব্যাপার। তবে দুটো symmetric matrix-এর গুণফল কিন্তু symmetric নাও হতে পারে। একই সমস্যা Hermitian, skew symmetric এবং skew Hermitian-দের বেলাতেও।

Exercise 22: এমন দুটো symmetric matrix দাও, যাদের গুণফল symmetric নয়। ■

এমন বহু symmetric matrix আছে যারা singular. কিন্তু যদি nonsingular হয়, তবে inverse-টাও symmetric হতে বাধ্য। এটা প্রমাণ করা খুব সহজ, খালি একটা জিনিস মনে রাখলেই হবে, যেটা আমরা একটু আগে শিখেছি--transpose-এর inverse, আর inverse-এর transpose একই জিনিস, অর্থাৎ

$$\left(A'\right)^{-1} = \left(A^{-1}\right)'.$$

এটা ব্যবহার করলেই নীচের অংকটা হয়ে যাবে।

Exercise 23: If A is a nonsingular, symmetric matrix, then show that A^{-1} is also symmetric. ■

Exercise 24: যদি A একটা nonsingular, Hermitian matrix হয়, তবে কি A^{-1}-ও Hermitian হতে বাধ্য? ■

Example 17: Let A be an $n \times n$ matrix whose elements are complex numbers. Show that $A + A^*$ is Hermitian, where $A^* = (\bar{A})^t$.[2] **(2003)**

SOLUTION: পরিচিত কয়েকটা ধর্ম ব্যবহার করে গেলেই অংকটা হয়ে যাবে। আমরা ধর্মগুলো সংক্ষেপে প্রমাণ করে দেব।

We know that

$$(A^*)^* = A. \qquad (*)$$

[[<u>Because:</u>

$$(A^*)^* = \left[\overline{[(\bar{A})^t]}\right]^t = \left[[(\bar{\bar{A}})^t]\right]^t = [[(A)^t]]^t = A.$$

]]

Also for any two matrices A, B of the same order we have

$$(A+B)^* = A^* + B^*. \qquad (**)$$

[[<u>Because:</u>

$$(A+B)^* = (\overline{A+B})^t = (\bar{A}+\bar{B})^t = (\bar{A})^t + (\bar{B})^t = A^* + B^*.$$

]]

So

$$\begin{aligned}(A+A^*)^* &= A^*+(A^*)^* \quad [\text{ by } (**)]\\ &= A^*+A \quad [\text{ by } (*)]\\ &= A+A^*.\end{aligned}$$

Hence $A+A^*$ is Hermitian, as required.

■

এই অংকটা থেকে শিক্ষণীয় এই যে, যেকোনো matrix থেকেই একটা Hermitian matrix বানানো যায়, খালি matrix-এর সঙ্গে তার conjugate transpose যোগ করে দিলেই হল। একইভাবে যেকোনো matrix থেকেই symmetric, skew symmetric বা skew Hermitian matrix-ও বানানো যায়। নীচের অংকটা তারই হদিশ দিয়েছে।

Exercise 25: For any $n\times n$ matrix A, show that

1. $A+A'$ is symmetric.
2. $A-A'$ is skew symmetric.
3. $A-A^*$ is skew Hermitian.
4. AA^* is Hermitian.
5. AA' is symmetric.

■

যেকোনো matrix দিয়েই যেমন symmetric, skew symmetric, Hermitian বা skew Hermitian matrix বানানো যায়, তেমনি এর উল্টোদিকটাও ঠিক--খালি symmetric আর skew symmetric matrix ব্যবহার করে যেকোনো square matrix বানানো যায়। একই কাজ করা যায় Hermitian আর skew Hermitian matrix-দের দিয়েও। কীভাবে, সেটাই এবারের অংকটার বক্তব্য।

Example 18: Show that any square matrix is expressible as a sum of a symmetric and a skew-symmetric matrix. Also show that this representation is unique.[2+1] **(2007)**

SOLUTION:

<u>Existence:</u>

Let A be any square matrix.

Then

$$A=\tfrac{1}{2}(A+A')+\tfrac{1}{2}(A-A').$$

Here $\frac{1}{2}(A+A')$ is a symmetric matrix because

$$[\tfrac{1}{2}(A+A')]'=\tfrac{1}{2}(A'+A)=\tfrac{1}{2}(A+A').$$

> Also, $\frac{1}{2}(A - A')$ is a skew symmetric matrix because
> $$[\tfrac{1}{2}(A - A')]' = \tfrac{1}{2}(A' - A) = -\tfrac{1}{2}(A - A').$$

এবার uniqueness দেখাব--

> Uniqueness:
> Let $A = P + Q$, where $P' = P$ and $Q' = -Q$.
> Then
> $$A + A' = P + Q + P' + Q' = P + Q + P - Q = 2P.$$
> So P must be $\frac{1}{2}(A + A')$.
> Hence Q must be $A - P = \frac{1}{2}(A - A')$.
> Thus, P, Q are uniquely determined by A.

■

কায়দাটা হাতেকলমে প্রয়োগ করে দেখার একটা সুযোগ দিয়েছে নীচের অংকটায়। অবশ্য এখানে আমরা কাজ করব Hermitian matrix নিয়ে।

Example 19: Express the matrix

$$A = \begin{bmatrix} 1 & 2+i & 1-i \\ 2-i & 1+2i & 3 \\ 2+i & 2 & 1+i \end{bmatrix}$$

as the sum of a Hermitian matrix and a skew Hermitian matrix.[2] **(2013.1bii)**
SOLUTION: প্রথমে কী করব লিখে নিই--

> For any square matrix A the matrix
> $$B = \tfrac{1}{2}(A + A^*)$$
> is Hermitian, and
> $$C = \tfrac{1}{2}(A - A^*)$$
> is skew Hermitian.
> Also $A = B + C$.

এবার হাতে কলমে--

> Here
> $$A^* = \begin{bmatrix} 1 & 2+i & 2-i \\ 2-i & 1-2i & 2 \\ 1+i & 3 & 1-i \end{bmatrix}$$

So

$$B = \begin{bmatrix} 1 & 2+i & \frac{3}{2}-i \\ 2-i & 1 & \frac{5}{2} \\ \frac{3}{2}+i & \frac{5}{2} & 1 \end{bmatrix}$$

and

$$C = \begin{bmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & 2i & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & i \end{bmatrix}.$$

■

যদি কিছু symmetric, skew symmetric, Hermitian ইত্যাদি matrix দেওয়া থাকে, তবে তাদের মিলিয়ে মিশিয়ে আরও নানারকম symmetric, Hermitian ইত্যাদি matrix বানানো যায়। এর মধ্যে আশ্চর্যের কিছু নেই। নীচের কয়েকটা অংক এইসব নিয়েই।

প্রথমে একটা বোকাবোকা প্রশ্ন দিয়ে শুরু করি, খুব চট্ করে উত্তর দাও দেখি--

Exercise 26: "যদি A একটা symmetric matrix হয়, তবে $-A$ অবশ্যই skew symmetric হবে"--ঠিক না ভুল? ■

Example 20: If A is a skew symmetric matrix then show that the matrix A^2 is symmetric. **(2002)**

SOLUTION:

Given that A is skew symmetric,

ie, $A' = -A$.

To show that A^2 is symmetric,

ie, $(A^2)' = A^2$.

Now

$$\begin{aligned} (A^2)' &= (AA)' \\ &= A'A' \\ &= (-A)(-A) \quad [\because A' = -A] \\ &= A^2, \end{aligned}$$

as required.

■

Example 21: If $S = P + iQ$ is a skew Hermitian matrix, then show that P is a real skew symmetric matrix and Q is a real symmetric matrix.[2] **(2006)**

SOLUTION: এই অংকটা ঠিক লেখেনি। এখানে আমাদের ধরে নিতে হবে যে P, Q দুজনেই real matrix.

We shall assume that P, Q are real matrices.
S is skew Hermitian, so $S^* = -S$.
Now

$$S^* = (P+iQ)^* = (\overline{P+iQ})' = (P-iQ)' \quad [\because P, Q \text{ real matrices}]$$
$$= P' - iQ'.$$

Hence $P' - iQ' = -P' - iQ'$.
So equating real and imaginary parts,

$$P' = -P \text{ and } -Q' = -Q.$$

Hence P is real skew symmetric and Q is real symmetric.

■

Exercise 27: If A is a Hermitian matrix, then show that iA is a skew-Hermitian matrix.[2] **(2008)** ■

Exercise 28: Show that B^*AB is Hermitian or skew Hermitian according as A is Hermitian or skew Hermitian, B being any matrix having the same order as the matrix A, and B^* being the transpose conjugate of B.[2] **(2012.9bii)** ■

Exercise 29: A and B are Hermitian matrices. Show that $AB + BA$ is Hermitian and $AB - BA$ is skew-Hermitian.[2] **(2007)** ■

এবার কিছু অংক দিই যেগুলো নিছক কসরতের অংক। Linear algebra শেখার জন্য এগুলো তেমন দরকারী কিছু নয়। প্রথম বার পড়ার সময়ে এগুলো বাদ দিয়ে যাওয়াই ভালো।

Exercise 30: Show that $(I_3 - A)(I_3 + A)$ is a symmetric matrix when A is a 3×3 symmetric matrix or a skew symmetric matrix.[3] **(2003)** ■

Exercise 31: A is a 3×3 non-null real matrix and $A^2 - A - I_3$ is a null matrix. Show that A^{-1} exists and $A^{-1} = A - I_3$.[2] **(2004)** ■

5.2 Orthogonal আর unitary matrix

আমরা কোনো matrix-এর transpose বার করা জানি। আরও শিখেছি যে, কোনো কোনো square matrix-এর inverse বার করা যায়। এক ধরণের matrix আছে, তাদের inverse আর transpose একই জিনিস! এদের বলে **orthogonal matrix**. সহজতম উদাহরণ হল I. এর inverse-ও I, আবার transpose-ও I.

Example 22: এই matrix-গুলোর মধ্যে কে কে orthogonal?

$$A=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},\quad B=\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{bmatrix},\quad C=\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}.$$

SOLUTION: আমাদের দেখতে হবে transpose আর inverse একই হয় কি না। তার জন্য একটা কায়দা হল transpose আর inverse বার করে মিলিয়ে দেখা। কিন্তু inverse বার করা মোটেই সহজ নয় (আমরা এখনও এই বইতে সেরকম কায়দা আলোচনাও করি নি)। তবে আমরা জানি যে, দুটো square matrix পরস্পরের inverse হবে যদি ওদেরকে গুণ করলে I পাওয়া যায়। সুতরাং একটা square matrix-কে তার transpose দিয়ে গুণ করলে যদি I আসে, তবেই বলা যাবে সেটা orthogonal. পরীক্ষা করে দেখা যাক। এখানে--

$$A'A=I,\quad B'B\neq I,\quad C'C=I.$$

তার মানে A আর C হল orthogonal, কিন্তু B নয়। ■

যেভাবে এই অংকটা করলাম সে ভাবেই orthogonal matrix-এর সংজ্ঞা লেখা যায়--

DEFINITION: Orthogonal matrix

A matrix A is called an **orthogonal matrix** if $AA'=A'A=I$.

বলাই বাহুল্য যে orthogonal matrix-রা square হতে বাধ্য, নইলে AA' আর $A'A$-এর সাইজ সমানই হত না। শুধু square-ই নয়, ওকে invertible-ও হতে হবে, কারণ A'-টাই তো A^{-1} হবে। Inverse-এর সংজ্ঞার পরে যেমন বলেছিলাম এখানেও তাই--যদি A একটা square matrix হয়, তবে $A'A=I$ আর $AA'=I$ এই দুটো শর্তের মধ্যে যে কোনো একটা পরীক্ষা করাই যথেষ্ট। একটা ঠিক হলে অন্যটাও ঠিক হবেই।
আমরা transpose আলোচনার সময়ে conjugate transpose-এর কথাও বলেছিলাম, যেখানে transpose নেওয়ার সাথে সাথে প্রত্যেকটা entry-র আবার conjugate-ও করা হয়। অবশ্য entry-রা সবাই real number হলে তাতে কিছুই এসে যাবে না, তাই সাধারণতঃ complex number-এর ক্ষেত্রেই ব্যাপারটা করা হয়। যদি আমরা orthogonal matrix-এর সংজ্ঞায় transpose-এর জায়গায় conjugate transpose লিখি, তবে পাব **unitary matrix**-এর সংজ্ঞা--

DEFINITION: Unitary matrix

A complex square matrix A is called a **unitary matrix** if $A^*A=AA^*=I$.

Example 23: দেখাও যে, $A=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ i & -i\end{bmatrix}$ একটা unitary matrix.

SOLUTION: এখানে $A^*=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -i\\ 1 & i\end{bmatrix}$ এবং $A^*A=AA^*=I$. তাই A একটা unitary matrix. ■

Orthogonal বা unitary matrix-দের দিয়ে কী হয়? তার খানিকটা উত্তর পাবে সপ্তম আর অষ্টম অধ্যায়ে গিয়ে। আপাততঃ নীচের কয়েকটা অংক করলে বুঝবে যে এদের নিয়ে কাজ করাটা সুবিধাজনক।

Exercise 32: দেখাও যে, A যদি একটা orthogonal matrix হয়, তবে A' এবং $-A$-ও তাই হবে। একইভাবে, A একটা unitary matrix হলে A^* এবং $-A$-ও তাই হবে। ■

Exercise 33: যদি A, B দুজনেই একই সাইজের দুটো orthogonal (বা unitary) matrix হয়, তবে দেখাও যে AB-ও তাই হবে। ■

এবার একটা সামান্য অন্যরকম অংক দিই, MCQ-এর আকারে।

Exercise 34: The number of orthogonal diagonal matrices of size $n \times n$ is
(A) 1 (B) n (C) 2^n (D) ${}^nC_2$.
HINT: একটু ধরিয়ে দিই। প্রথমে $n = 2$ নিয়ে ভাবো। যেকোনো 2×2 সাইজের diagonal matrix-ই দেখতে $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$-এর মত। আমাদের দরকার $A'A = I$. দেখো তো এর থেকে a, b-এর উপরে কী কী শর্ত পাও। ■

Example 24: Determine the values of α, β, γ where

$$\begin{bmatrix} 0 & 2\beta & \gamma \\ \alpha & \beta & -\gamma \\ \alpha & -\beta & \gamma \end{bmatrix}$$

is orthogonal.[3] **(2008)**
SOLUTION: একটা matrix A-কে orthogonal হতে হলে $A'A = I$ হতে হবে।

Let

$$A = \begin{bmatrix} 0 & 2\beta & \gamma \\ \alpha & \beta & -\gamma \\ \alpha & -\beta & \gamma \end{bmatrix}$$

Thus we need

$$A'A = \begin{bmatrix} 2\alpha^2 & 0 & 0 \\ 0 & 6\beta^2 & 0 \\ 0 & 0 & 3\gamma^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

ie,

$$2\alpha^2 = 6\beta^2 = 3\gamma^2 = 1.$$

So

$$\alpha = \pm\frac{1}{\sqrt{2}}, \quad \beta = \pm\frac{1}{\sqrt{6}}, \quad \gamma = \pm\frac{1}{\sqrt{3}}.$$

এখানে $AA' = I$ নিয়েও কাজ করা যেত, যদিও তাতে সমাধানটা সামান্য বেশী লম্বা হত (করেই দেখো!)। ■

নীচের তিনটে অংক নিছক কসরতের অংক। Linear algebra শেখার জন্য একেবারেই দরকারী কিছু নয়। প্রথম বার পড়ার সময়ে বাদ দিয়ে যাওয়াই ভালো।

Exercise 35: If $(I + A)^{-1}(I - A)$ is a skew-symmetric matrix, then show that A is a real orthogonal matrix.[3] **(2012.9bi)**

HINT:
এখানে ধরে নিতে হবে যে, A একটা real matrix, মানে এর সবগুলো entry-ই real number. ■

এই অংকটারই কতকটা উল্টোদিক রয়েছে নীচের অংকে--

Exercise 36: A is a 3×3 real matrix and $(I_3 - A)(I_3 + A)^{-1}$ is orthogonal. Show that A is a skew-symmetric matrix.[3] **(2007)** ■

Exercise 37: If A is a real orthogonal matrix and $I + A$ is nonsingular, prove that $(I+A)^{-1}(I-A)$ is skew symmetric, where I denotes the identity matrix of suitable order.[3] **(2009)** ■

5.3 Idempotent matrix

ধরো এমন একটা matrix দিলাম A, যার জন্য $A^2 = A$ হয়। যেমন

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

তবে সেরকম A-কে বলব **idempotent**. (A^2 বার করে দেখে নিও সত্যিই এখানে $A^2 = A$ হচ্ছে কিনা!) আশা করি সন্দেহ নেই যে idempotent matrix মাত্রেই square হতে বাধ্য। তা নইলে A^2 বার করাই যেত না! Idempotent matrix-রা কেন গুরুত্বপূর্ণ সেটা চট্ করে এখনই বোঝা মুস্কিল। এমনি দেখলে মনে হয় যে, এদের কোনোই গুণ নেই। দুটো idempotent matrix-এর যোগফল, বিয়োগফল বা গুণফল কোনোটাই idempotent নাও হতে পারে। অবশ্য একটা idempotent matrix-এর transpose সর্বদা idempotent-ই হবে। তার উপরে idempotent matrix-রা সাধারণতঃ nonsingular-ও হয় না। কেন হয় না বুঝতেই পারছ-- যদি $A^2 = A$ হয় আর A^{-1} থাকে, তবে দু দিককে A^{-1} দিয়ে গুণ করে দিলেই $A = I$ পাবে। অবশ্যই I একটা idempotent matrix, কিন্তু ওইটাই একমাত্র nonsingular, idempotent matrix.

কোনো গুণই যদি না থাকে, তবে খামোখা idempotent matrix-দের কথা শিখব কেন? কারণ আসলে এরা খুবই কাজের জিনিস। সেটা বুঝতে পারবে পরে, যখন projection এবং linear transformation বলে দুটো জিনিস শিখবে।

Exercise 38: যদি A একটা idempotent matrix হয় তবে দেখাও যে, যেকোনো $n \in \mathbb{N}$-এর জন্যই $A^n = A$ হবে। ■

Example 25: Prove that if A, B are two matrices such that $AB = A$ and $BA = B$ then A, B are idempotent.[3] **(2001)**

SOLUTION:

$$\begin{aligned} A^2 &= (AB)(AB) \quad [\because A = AB.] \\ &= A(BA)B \\ &= ABB \quad [\because B = BA.] \\ &= AB \quad [\because A = AB.] \\ &= A \quad [\because A = AB.]. \end{aligned}$$

Hence A is idempotent.

Similarly,

$$\begin{aligned} B^2 &= (BA)(BA) \quad [\because B = BA.] \\ &= B(AB)A \\ &= BAA \quad [\because A = AB.] \\ &= BA \quad [\because B = BA.] \\ &= B \quad [\because B = BA.]. \end{aligned}$$

Hence B is idempotent.

■

DAY 6
Field

6.1 কোন্ ধরণের সংখ্যা?

আমরা এতক্ষণ বলে আসছি যে, একটা matrix-এর entry-গুলো একেকটা "সংখ্যা"। কথা হল, "সংখ্যা" তো নানারকম হয়--

- integer (পূর্ণসংখ্যা, $\mathbb{Z}$), যেমন $0, \pm1, \pm2, \pm3, ...$
- rational (মূলদ সংখ্যা, $\mathbb{Q}$) , যাদের মধ্যে integer-রা তো আছেই, সেই সঙ্গে আছে fraction-রা, যেমন $\frac{1}{2}$, $\frac{34}{5}$, $-\frac{65}{1000}, ...$।
- real number (বাস্তব সংখ্যা, $\mathbb{R}$), যাদের মধ্যে rational-রা তো আছেই, সেই সঙ্গে আছে $\sqrt{2}$, π ইত্যাদিরা, এবং আরো এমন সব সংখ্যা যাদের কোনো ভদ্র নাম পর্যন্ত নেই।
- complex number (জটিল সংখ্যা, $\mathbb{C}$), যেগুলো দেখতে $a + ib$-এর মত, যেখান $a, b \in \mathbb{R}$ আর i একটা অদ্ভুত জিনিস যাকে square করলে -1 হয়!

এর মধ্যে কোন্ ধরণের সংখ্যা নিয়ে আমরা কাজ করব? নাকি এর বাইরেও আরও নতুন কোনো ধরণের সংখ্যা সম্ভব? হ্যাঁ, এর বাইরেও আরও বহু ধরণের সংখ্যা সম্ভব, যেগুলো দৈনন্দিন জীবনে চট্ করে চোখে না পড়লেও অনেক জায়গাতেই ব্যবহার হয়। যেমন ধরো তোমার মোবাইলের সিগনালগুলো সংখ্যার মাধ্যমেই যায়, এবং সেগুলো আমাদের পরিচিত সংখ্যা নয়। DVD-র গায় লেজার ফেলে অসংখ্য ফুটো ফুটো করে গান, সিনেমা ইত্যাদি সংখ্যা দিয়ে খোদাই করে লেখা থাকে। এই সংখ্যারাও একটা নতুন ধরণের সংখ্যা। আমাদের পরিচিত real আর complex সংখ্যারা তো বটেই, এমন কি এইসব অদ্ভুত সংখ্যা দিয়েও একটা matrix তৈরী হতে পারে। ঠিক কোন্ ধরণের সংখ্যা linear algebra-য় ব্যবহার করা, সেটা একটু বুঝে নেওয়া যাক।

লক্ষ কর আমরা যখন matrix নিয়ে কাজ করি, তখন তার entry-দের নিয়ে যোগ-বিয়োগ-গুণ-ভাগ করি। সুতরাং সংখ্যাগুলো এমন হতে হবে যেন ওদেরকে নিশ্চিন্তে যোগ-বিয়োগ-গুণ-ভাগ করা যায়। ফলে "খালি integer-দের নিয়ে কাজ করব, অন্য কোনো সংখ্যাকে ঢুকতে দেব না"-জাতীয় গোঁ ধরে থাকলে মুস্কিল। এতে যোগ-বিয়োগ-গুণে অসুবিধা হবে না, কিন্তু ভাগ করতে অসুবিধা হবে, কারণ একটা integer-কে আরেকটা integer দিয়ে ভাগ করলে উত্তরটা integer নাও হতে পারে, যেমন 2-কে 3 দিয়ে ভাগ করলে $\frac{2}{3}$ হয়, যেটা integer নয়। অবশ্য যদি একটু উদার হয়ে rational-দের ব্যবহার করতে রাজি থাকি, তবে কোনো সমস্যা নেই। যদি তুমি যাবতীয় real number নিয়ে কাজ করতে রাজি থাকো, তবে তো আরও ভালো।

যদি complex numberদের নিয়ে কাজ কর, সেখানেও যোগ-বিয়োগ-গুণ-ভাগের সব নিয়মকানুনই সাধারণ সংখ্যাদের মতই। ফলে সেখানেও অসুবিধা হবে না। আরও যেসব অদ্ভুত সংখ্যাদের কথা বললাম, তাদের বেলাতেও যদি যোগ-বিয়োগ-গুণ-ভাগের নিয়ম আমাদের পরিচিত নিয়মগুলোর মতই হয়, তবে তাদের নিয়েও চমৎকার linear algebra করা যাবে। অতএব আমরা এক কথায় বলতে পারি--

> এরকম যেকোনো সংখ্যাদের নিয়েই linear algebra করতে পারি, যাদেরকে আমাদের পরিচিত নিয়মে যোগ-বিয়োগ-গুণ-ভাগ করা যায়।

"পরিচিত নিয়ম" মানে যে নিয়মগুলো ছোটোবেলায় স্কুলে শিখি আর কি, যেমন $a(b+c) = ab+ac$, এই সব। এই রকম নিয়ম-মেনে-চলা সংখ্যাদের set-দের একটা নাম আছে--field. এদের সংজ্ঞার মধ্যেই বিভিন্ন "পরিচিত নিয়ম"-এর একটা তালিকা পাবে।

DEFINITION: Field

By a **field** we mean a set F along with two binary operations, addition $(+)$ and multiplication $(\times)$, such that

1. $\forall a, b \in F \quad a+b \in F$ and $a \times b \in F$.
2. $+$ and $\times$ are both associative, *i.e.*,
$$\forall a, b, c \in F \quad (a+b)+c = a+(b+c), \quad (a \times b) \times c = a \times (b \times c).$$
3. $+$ and $\times$ are both commutative, *i.e.*,
$$\forall a, b \in F \quad a+b = b+a, \quad a \times b = b \times a.$$
4. $+$ has an identity element 0 (**additive identity**) and $\times$ has an identity element $1(\neq 0)$, (**multiplicative identity**) *i.e.*,
$$\exists 0 =\neq 1 \in F \quad \forall a \in F \quad a+0 = a, \quad 1a = a.$$
5. every element has an additive inverse (called its negative), *i.e.*,
$$\forall a \in F \quad \exists b \in F \quad a+b = 0.$$
This b is written as $-a$.
6. every nonzero element has a multiplicative inverse, *i.e.*,
$$\forall a(\neq 0) \in F \quad \exists b \in F \quad ab = 1.$$
This b is written as a^{-1} or $\frac{1}{a}$.
7. $\times$ distributes over $+$, *i.e.*,
$$\forall a, b, c \in F \quad a \times (b+c) = a \times b + a \times c.$$

এই সংজ্ঞাটা সম্পর্কে দুটো কথা--

- এখানে binary operation মানে হল যোগ, বিয়োগ গুণের মত এমন কিছু কাজ যা দিয়ে F-এর দুটো element-কে মিলিয়ে F-এরই আরেকটা element পাওয়া যাবে। অংকের ভাষায় বললে একটা binary operation হল $F \times F$ থেকে F-এ একটা function.
- আমরা যখন $a - b$ লিখি, তার মানে আসলে $a + (-b)$. একইভাবে $\frac{a}{b}$ মানে হল $a \times b^{-1}$.

Field-এর অতি পরিচিত কিছু উদাহরণ হল $\mathbb{Q}$, $\mathbb{R}, \mathbb{C}$. কিন্তু $\mathbb{Z}$ একটা field নয়, কারণ একটা $2 \in \mathbb{Z}$ কিন্তু $\frac{1}{2} \notin \mathbb{Z}$. এই কয়টা পরিচিত field ছাড়াও আরও নানারকম field হয়, যাদের খানিকটা স্বাদ এবার দেব (আমাদের বাংলায়-বোঝানো-ইংরাজি-বই সিরিজের Modern Algebra বইতে আরও তথ্য পাবে)। $\mathbb{R}$ ইত্যাদিরা হল যাকে বলে infinite field, কারণ ওদের মধ্যে অসংখ্য element আছে। অনেক field কিন্তু finite হয়, যেমন নীচের উদাহরণে।

Example 26: এই সেটটা নাও, $F = \{0, 1, 2\}$. এই সেটের উপর যোগ, বিয়োগ, গুণ করার করার কায়দা হল প্রথমে

এমনিভাবে যোগ, বিয়োগ, গুণ করে শেষে 3 দিয়ে ভাগ করে remainder (ভাগশেষ) নিতে হবে। এরকম operation-দের বলে **modulo 3 operation.** যাতে সাধারণ যোগের সাথে গুলিয়ে না যায়, তাই modulo 3 addition-কে অনেক সময়ে $\oplus_3$ লেখা হয়। একইভাবে $\otimes_3$ এবং $\ominus_3$ চিহ্নগুলোও ব্যবহৃত হয়। যেমন যদি বলি $2 \otimes_3 (2 \oplus_3 2 \ominus_3 1)$ করতে তবে তুমি এইভাবে এগোবে--প্রথমে এমনি যোগ-বিয়োগ-গুণ দিয়ে অংকটা করবে, তাতে আসবে

$$2 \times (2 + 2 - 1) = 6.$$

এবার একে 3 দিয়ে ভাগ করলে remainder হবে 0. সুতরাং সব মিলিয়ে পেলাম

$$2 \otimes_3 (2 \oplus_3 2 \ominus_3 1) = 0.$$

ভারি অদ্ভুত, সন্দেহ নেই! এভাবে না হয় যোগ, বিয়োগ, গুণের ব্যবস্থা হল, কিন্তু ভাগের বেলায় তো আর এই কায়দা খাটবে না। যেমন যদি বলি 1-কে 2 দিয়ে ভাগ করতে, তবে এমনিভাবে এগোলে $\frac{1}{2}$ আসবে, যেটা আবার একটা fraction, তাকে আবার 3 দিয়ে ভাগ করে remainder কী বার করব? তাই ভাগের জন্য নিয়মটা অন্য। আমরা দেখব 2-কে F-এর কোন্ element দিয়ে গুণ করলে 1 আসে। বিভিন্ন সংখ্যা দিয়ে 2-কে গুণ করে দেখা যাক--

$$\begin{aligned} 0 \otimes_3 2 &= 0 \text{ (নাঃ, 1 হল না!)} \\ 1 \otimes_3 2 &= 2 \text{ (এবারও 1 হল না!)} \\ 2 \otimes_3 2 &= 1 \text{ (পেয়েছি!!)} \end{aligned}$$

সুতরাং আমরা বলব যে আমাদের নতুন নিয়মে 1-কে 2 দিয়ে ভাগ করলে 2 হয়, মানে $1 \oplus\!\!\!\!\div_3 2 = 2$. এরকম আজব নিয়ম দেখে তুমি হয়তো খাবি খাচ্ছ, তাই বোঝার সুবিধার জন্য ভাগের এই নতুন নিয়মের পুরো নামতাটাই দিলাম নীচে--

$$0 \,\bigcirc\!\!\!\!\!\div_3\, 1 = 0 \,\bigcirc\!\!\!\!\!\div_3\, 2 = 0, \quad 1 \,\bigcirc\!\!\!\!\!\div_3\, 1 = 1, \quad 1 \,\bigcirc\!\!\!\!\!\div_3\, 2 = 2, \quad 2 \,\bigcirc\!\!\!\!\!\div_3\, 1 = 2, \quad 2 \,\bigcirc\!\!\!\!\!\div_3\, 2 = 1.$$

লক্ষ কর যে সাধারণ সংখ্যাদের মত এখানেও 0 দিয়ে ভাগ করা চলে না। এটা যে একটা field, সেটা চাইলে সংজ্ঞা ধরে ধরে প্রমাণ করে দেখতে পারো। এই field-টাকে বলে $\mathbb{Z}_3$. যেহেতু এটা একটা field , তাই স্কুলে শেখা যোগ, বিয়োগ, গুণ, ভাগের যাবতীয় ফর্মুলা ইত্যাদি এখানেও খাটবে। যেমন ধরো এখানেও $(a + b)^2 = a^2 + b^2 + 2ab$ হবে। এখানেও

$$1 + a + \cdots + a^n = \frac{a^{n+1} - a}{a - 1} \quad (a \neq 1)$$

হবে, ইত্যাদি!!! অথচ আমাদের field-এ সংখ্যা আছে মাত্র তিনটে, তাই নিয়েই স্কুলে যোগ, বিয়োগ, গুণ, ভাগের যতরকম কায়দা শিখেছ সব করা যাচ্ছে! ■

Exercise 39: একইভাবে দেখাও যে $\mathbb{Z}_2 = \{0, 1\}$ নিয়ে যদি আমরা modulo 2 যোগ, বিয়োগ, গুণ করি, তবে $\mathbb{Z}_2$-ও একটা field হবে। এটাই বিশ্বের সবচেয়ে ছোটো field. এখানে $1 \oplus_2 (1 \oplus_2 1 \oplus_2 1)$ কত হবে? ■

তবে ভেবে বোসো না যেন যে, এরকম modulo operation দিয়ে সবসময়ে পার পাওয়া যাবে। যেমন নীচের অংকটায়।

Example 27: যদি আমাদের set-টা নিই $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ আর operation-গুলো modulo 4 করি, তবে $1 \oplus_4 2$ কী হবে?

SOLUTION: এখানেও আমরা 2-কে বিভিন্ন সংখ্যা দিয়ে modulo 4 গুণ করে করে খুঁজব কখন 1 আসে--

$$\begin{aligned} 2 \otimes_4 0 &= 0 \text{ (নাঃ, 1হল না} \\ 2 \otimes_4 1 &= 2 \text{ (এবারেও 1 হল না)} \\ 2 \otimes_4 2 &= 0 \text{ (এবারও হল না)} \\ 2 \otimes_4 3 &= 2 \text{ (শেষ আশাও গেল!)} \end{aligned}$$

যেহেতু কোথাওই 1 পেলাম না, তাই এখানে field তৈরী হল না, কারণ field হবার জন্য সব nonzero element-এরই multiplicative inverse থাকা দরকার। ■

এই অংকটা দেখে ভেবো না যেন, চার সাইজের কোনো field হয় না। সেটা অবশ্যই হয়, কিন্তু সেখানে operation-গুলো অন্যভাবে নিতে হয়। নীচের অংকটা সেটা নিয়েই।

Example 28: আবার আমরা চার সাইজের একটা set-টা নেব, কিন্তু element-গুলোর নাম দেব একটু অন্যভাবে, $\{0, 1, x, 1+x\}$. যোগ, বিয়োগ, গুণের নিয়মও অন্যরকম-- প্রথমে সাধারণ polynomial-দের মতই করব, খালি কোথাও x^2 এসে গেলে তার জায়গায় $x+1$ বসিয়ে দেব। সব শেষে coefficient-গুলোকে 2 দিয়ে ভাগ করে, খালি remainder-টুকু রাখব। যেমন $x+(1+x)$, বার করার জন্য প্রথমে এমনিভাবে এগিয়ে পাব $2x+1$. এবার 2 দিয়ে ভাগ করে remainder নিলে পাব 1. একইভাবে, $(1+x)^2$ যদি এমনি বার করতাম তবে পেতাম $1+2x+x^2$. এখানে x^2-এর জায়গায় $1+x$ বসালে $1+2x+1+x = 2+3x$. এবার modulo 2 করলে হবে x. ■

কোথা থেকে এই আজব নিয়ম এল, সেটা বোঝানোর জায়গা এটা নয়। বিভিন্ন সাইজের field বানানোর জন্য এরকম উপযুক্ত binary operation বানানোর একটা চমৎকার কায়দা আছে। সেটা লাগিয়েই এই আজব নিয়মটা পাওয়া গেছে।
কোন্ কোন্ সাইজের finite field সম্ভব, সে বিষয়ে একটা theorem আছে--

THEOREM

Every finite field has size p^n for some prime p and some $n \in \mathbb{N}$. Conversely, for any prime p and any $n \in \mathbb{N}$ there is exactly one field of size p^n.

এই theorem-এর প্রমাণ মোটেই সহজ নয়। যারা field extension নামের একটা জিনিস শিখবে, তারা এ সম্পর্কে জানতে পারবে। p^n সাইজের field-রা বাদে বাকি সব field-ই infinite সাইজের। এই প্রসঙ্গে একটা জিনিস লক্ষ করো। যখন আমরা $\mathbb{Q}$ বা $\mathbb{R}$ বা $\mathbb{C}$ নিয়ে কাজ করি, তখন পরপর 1 যোগ করে গেলে $1,\ 1+1,\ 1+1+1$, এইভাবে চলতেই থাকে, চলতেই থাকে, কোনোদিনই শূন্য হয়ে যায় না। কিন্তু আমরা যখন $\mathbb{Z}_3$-এর মধ্যে কাজ করছিলাম, তখন $1 \oplus_3 1 \oplus_3 1 = 0$ হয়েছিল। আবার যখন চার সাইজের field-টায় কাজ করছিলাম, তখন $1 \oplus 1 = 0$ হয়েছিল। এইভাবে যেকোনো finite field-এই দেখা যায় যে, কয়েকটা 1 পরপর যোগ করে গেলে এক সময়ে শূন্য হয়ে যায়। সবচেয়ে কম যতগুলো 1 যোগ করলে শূন্য

পাওয়া যায়, সেই সংখ্যাটাকে বলে সেই field-এর **characteristic.** যেমন তিন সাইজের field-টার characteristic ছিল 3 আর চার সাইজেরটার characteristic ছিল 2. যদি field-টা infinite হয়, তবে যতবারই 1 যোগ করো না কেন, কোনোদিনই শূন্য বানাতে পারবে না। এরকম field-এর characteristic হল 0.

DEFINITION: Characteristic

The characteristic of a field is the smallest $n \in \mathbb{N}$ such that

$$\underbrace{1 + \cdots + 1}_{n} = 0$$

in that field. If no such the n exists, then we say that the field has characteristic 0.

আমরা আগেই দেখেছি যে, field-দের সাইজ হয় infinite হয়, আর নয়তো p^n জাতীয় হয়, যেখানে p হল কোনো prime এবং $n \in \mathbb{N}$. এই সাইজটুকু জানলেই কোনো field-এর characteristic বলে দেওয়া যায়--

THEOREM

A finite field of size p^n (where p is a prime and $n \in \mathbb{N}$) has characteristic p. All infinite fields have characteristic ∞.

আমরা এ বইতে field হিসেবে প্রায় সর্বত্রই $\mathbb{R}$ বা $\mathbb{C}$-ই ব্যবহার করব। যে সব আলোচনায় খালি যোগ-বিয়োগ-গুণ-ভাগ লাগে, সেগুলো যেকোনো field-এর ক্ষেত্রেই খাটবে। যেমন নীচের উদাহরণে--

Example 29: আমরা এই অধ্যায়ের গোড়াতেই শিখেছিলাম যে $ad - bc \neq 0$ হলে

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

হয়। এটা প্রমাণ করার জন্য খালি যোগ-বিয়োগ-গুণ-ভাগ করতে পারাই যথেষ্ট, তাই এই জিনিসটা যেকোনো field-এর বেলাতেই খাটে। এই কায়দায় $\mathbb{Z}_3$-র উপরে এই matrix-টার inverse বার করো--

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}.$$

SOLUTION:এখানে $a = b = 1, c = 0, d = 2$. তাই $a \otimes_3 d \ominus_3 b \otimes_3 c = 2 \neq 0$. আমরা আগেই দেখেছি যে $1 \oslash_3 2 = 2$. তাই

$$A^{-1} = 2 \otimes_3 \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}.$$

এই শেষের ধাপে আমরা ব্যবহার করেছি যে -1 মানে আসলে 2 (কারণ $1 \oplus_3 2 = 0$)!
এবার এই A^{-1} দিয়ে A-কে গুণ করে দেখো, সত্যিই I পাবে! তার মানে আমাদের field-এ যোগ, বিয়োগ, গুণ, ভাগের নিয়ম যতই তুঘলকী হোক না কেন, matrix-এর inverse বার করার এই কায়দাটা একইভাবে খাটছে! ■

Exercise 40: $\mathbb{Z}_3$-র উপরে এই matrix-টার inverse বার করো--

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

■

Exercise 41: $\mathbb{Z}_2$-কে field হিসেবে নিয়ে AB বার করো, যেখানে--

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ আর } B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

■

6.2 যেখানে বিশেষ field দরকার

এ বইতে অধিকাংশ আলোচনাই যেকোনো general field-এর জন্য প্রযোজ্য হলেও কোনো কোনো ক্ষেত্রে কিছু বিশেষ ধরণের field-ই লাগবে। এবার সেরকম কিছু উদাহরণ দেখব।

আমরা যখন $\mathbb{R}$ নিয়ে কাজ করি, তখন যে খালি যোগ-বিয়োগ-গুণ-ভাগই করি, তা নয়। অনেক সময়ে দুটো সংখ্যার মধ্যে ছোটো-বড়ো তুলনাও করি, যেমন $2 < 3$. সব field-এর বেলায় যে এরকম তুলনা করা যাবে, এমন নয়। যেমন ধরো $\mathbb{Z}_3$-এর বেলায় -2 আর 1 আসলে একই জিনিস। এবার ভাবো $1 > 0$ বলবে, নাকি $1 < 0$ বলবে? এইরকম নানা অসুবিধার কারণে সব field-এর বেলায় $>$ বা $<$ বা $\leq$ বা $\geq$ ব্যবহার করা যায় না। সুতরাং যদি আমরা linear algebra-র কোনো আলোচনায় এই চিহ্নদের ব্যবহার করি, তবে সেই সব আলোচনাগুলো সব field-এর বেলায় প্রযোজ্য হবে না।

Example 30: যদি A কোনো 3×1 সাইজের matrix হয়, তবে $A'A$ হবে একটা 1×1 সাইজের matrix, মানে একটা সংখ্যা। এই সংখ্যাটা সর্বদাই ≥ 0 হবে। এ কথাটা কি ঠিক, নাকি ভুল?

SOLUTION: ধরো $A = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. তবে $A'A = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a^2 + b^2 + c^2$. যেহেতু এটা কয়েকটা square-এর যোগফল, তাই ফস্ করে মনে হতে পারে যে, এটা ≥ 0 হতে বাধ্য। কিন্তু সাবধান, এর জন্য আমরা মনে মনে কল্পনা করে ফেলেছি যে $a, b, c \in \mathbb{R}$. যদি অন্য কোনো field হয়, তবে সেখানে "$\geq$"-র কোনো অর্থ নাও থাকতে পারে। তাই বাক্যটা $\mathbb{R}$-এর বেলায় ঠিক হলেও, general কোনো field-এর বেলায় অর্থহীন। ■

এই উদাহরণটায় একটা জিনিস লক্ষ করার মত। সেটা হল "সংখ্যা" শব্দটার ব্যবহার। শব্দটা শোনামাত্র আমাদের মনে real সংখ্যাদের কথাই আসে, কারণ সেটাই আমাদের ছোটোবেলার অভ্যাস। কিন্তু linear algebra-র বেলায় "সংখ্যা" বলতে বোঝানো হচ্ছে "যে field-এর মধ্যে কাজ হচ্ছে, তার কোনো element"। যাতে মাথা গুলিয়ে না যায়, তাই সংখ্যা না বলে আমরা এদের বলি **scalar.**

এবার আরেকটা সমস্যার কথা বলি, যেখানে field-এর উপস্থিতিটা ভুলে গেলে বিপদ হয়।

Example 31: ধরো a, b এমন দুটো scalar যাতে $(a+b)^2 = a^2 + b^2$ হয়। এ থেকে কি তুমি জোর দিয়ে বলতে পারো যে, a আর b-র মধ্যে অন্ততঃ একজন শূন্য হবেই?

SOLUTION: এখানে $(a+b)^2 = a^2 + b^2$ বলেছে, তার মানে $a^2 + b^2 + 2ab = a^2 + b^2$. সুতরাং $2ab = 0$. সুতরাং $ab = 0$, তার মানে a আর b-এর মধ্যে একজন শূন্য হবেই।

সাবধান! এটা কি ঠিক বললাম? না, এখানে $2ab = 0$ থেকে আমরা ধাঁ করে $ab = 0$ লিখে দিতে পারি না। যদি field-এর characteristic হয় 2, তবে $2 = 1 + 1 = 0$ (অ্যাঁ?) তাই 2-টাকে কাটা যাবে না। বস্তুতঃ characteristic যদি 2 হয়, তবে $(a+b)^2 = a^2 + b^2$ হলেও $ab \neq 0$ হতে পারে, যেমন $a = b = 1$ নিলে--

$$(a+b)^2 = (1+1)^2 = 0$$

এবং

$$a^2+b^2 = 1^2+1^2 = 1+1 = 0.$$

এখানে যেটা বলা উচিত ছিল সেটা হল--

> যদি field-এর characteristic-টা 2 না হয়, তবে $(a+b)^2 = a^2+b^2$ হলে a বা b শূন্য হবেই।

■

এইটুকু পড়ে হয়তো তোমার অস্বস্তি হতে শুরু করেছে। প্রতি পদে পদে এরকম characteristic হিসেব করে এগোতে হলে তো মাথা ধরে যাবে! না, ব্যাপারটা অত খারাপ কিছু নয়। এরকম জিনিস অল্পই আসে। বেশীরভাগ ক্ষেত্রেই তুমি $\mathbb{R}$-এর উপর কাজ করছ মনে করেই এগোতে পারো, তাতে যেকোনো field-এর বেলাতেই ঠিক উত্তরই আসবে। যদি কোথাও তা না হয়, তবে এ বইতে তোমাকে আগেভাগেই সাবধান করে দেওয়া থাকবে।
এবার তৃতীয় একটা ক্ষেত্র আলোচনা করি। কোনো কোনো ক্ষেত্রে আমাদেরকে polynomial-দের zero বার করতে হবে, মানে polynomial-টা $p(x)$ হলে x-এর এমন value বার করতে হবে, যাতে $p(x)=0$ হয়। যেমন, x^2-2-এর দুটো zero, একটা হল $\sqrt{2}$ অন্যটা $-\sqrt{2}$. সব field-এর বেলায় সব polynomial-এর আবার zero পাওয়া যায় না। যেমন $\mathbb{Z}_3$-এর মধ্যে x^2-2-এর কোনো zero নেই (অর্থাৎ $x=0,1,2$ বসালে কখনোই $\mathbb{Z}_3$-এর মধ্যে $x^2-2=0$ হবে না, চেষ্টা করেই দেখো!)। আরেকটা উদাহরণ হল $\mathbb{R}$-এর মধ্যে x^2+1-এর কোনো zero নেই। কিছু কিছু field আছে, তাদের বেলায় সব polynomial-এরই zero থাকে, এরকম সুন্দর field-দের বলে **algebraically closed**. এরকম field-এর সবচেয়ে পরিচিত উদাহরণ হল $\mathbb{C}$. তাই linear algebra-র যেসব অংকে কোনো polynomial-এর zero বার করার দরকার পড়বে, সে অংকগুলো আমরা সাধারণতঃ $\mathbb{C}$ বা অন্য কোনো algebraically closed field-এর উপরে করব। অবশ্য সপ্তম অধ্যায়ের আগে সে প্রয়োজন হবে না।

Answers

1. Square matrix হলে। **2.** (A), (B), (D) সঠিক। (C) ভুল। **3.** অসম্ভব, যদি A-র (1,1)-th entry হয় a_{11}, তবে $A+A^*$-এর (1,1)-th entry হবে $a_{11}+\overline{a_{11}} = 2\mathrm{Re}(a_{11})$, যেটা একটা real number হতে বাধ্য।
4. $AB' = \begin{bmatrix} 30 & 8 \\ 40 & 10 \end{bmatrix}$, $A'B = \begin{bmatrix} 0 & 4 & 20 \\ 0 & 6 & 28 \\ 0 & 8 & 34 \end{bmatrix}$, AB বার করা অসম্ভব, কারণ A-এর column-সংখ্যা আর B-এর row-সংখ্যা অসমান, $(AA')^2 = \begin{bmatrix} 2690 & 4085 \\ 4085 & 6205 \end{bmatrix}$. **5.** ওর মানে হল প্রত্যেকের মোট কতজন করে friend আছে, যেমন প্রথম diagonal entry হল 3, অর্থাৎ আলমের মোট 3 জন বন্ধু আছে। **6.** এখানে
$F^2 = \begin{bmatrix} 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 \end{bmatrix}$. আমরা খালি এর off-diagonal entry-গুলো রাখলেই mutual friend-দের সংখ্যা
পেয়ে যাব-- $\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$. **7.** $c_{23} = \sum_{k=1}^4 a_{2k}b_{k3}$. **8.** $d_{21} = \sum_{i=1}^4 \sum_{j=1}^{10} a_{2i}b_{ij}c_{j1}$. **9.** না থেকে যাবেটা কোথায়? **10.** খালি দ্বিতীয় column-টা। সেটা হবে $\begin{bmatrix} 8 \\ 23 \end{bmatrix}$. **11.** $a=-2,\ b=1,\ c=\frac{3}{2}$, আর $d=-\frac{1}{2}$. **12.** না, A^{-1} থাকবে না, কারণ যাই $B_{2\times 2}$ নেব, $AB=O$ হয়ে যাবে। **13.** না, A^{-1} থাকবে না, কারণ যাই $B_{2\times 2}$ নেব, AB-র দ্বিতীয় row-টা $(0,0)$ হয়ে যাবে। **14.** $A^*\left(A^{-1}\right)^* = \left(A^{-1}A\right)^* = I^* = I$. আবার $\left(A^{-1}\right)^*A^* = \left(AA^{-1}\right)^* = I^* = I$. **15.** $\vec{\mathbb{1}}'A = K\vec{\mathbb{1}}'$, তাই $\vec{\mathbb{1}}' = K\vec{\mathbb{1}}'A^{-1}$. এখানে $K=0$ হতে পারে না,

কারণ বাঁদিকে $\mathbb{1} \neq \vec{0}$. অতএব $\frac{1}{K}\mathbb{1}' = \mathbb{1}'A^{-1}$. **16.** $\sum_k a_{ik}b_{kj}$-এ যখনই $k \neq i$ হবে অমনি $a_{ik} = 0$ হয়ে যাবে। সুতরাং খালি $k = i$ কেসটাই পড়ে থাকবে। **17.** $\sum_k a_{ik}b_{kj} = a_{ij}b_{jj}$ হবে কারণ এখন B-কে diagonal নিয়েছি, তাই $k \neq j$ হলে $b_{kj} = 0$ হবে। **18.** $c_{ij} = \sum_k a_{ik}b_{kj}$. যদি $i < j$ হয়, তবে $c_{ij} = 0$ দেখাব। Sum-টার মধ্যে যখন $k < j$, তখন $b_{kj} = 0$. আবার $k \geq j$ হলে $i < k$ হবে। তখন $a_{ik} = 0$. সুতরাং sum-টার প্রত্যেকটা term-ই শূন্য। তার মানে $i < j$ হলে $c_{ij} = 0$. **19.** $\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ \frac{5}{2} & -1 & 0 \\ \frac{-3}{10} & 0 & \frac{1}{10} \end{bmatrix}$. **20.** $a_{ii} = 0$ হবে, কারণ $a_{ii} = -a_{ii}$.

21. $Re(a_{ii}) = 0$ হবে, কারণ $a_{ii} = -\overline{a_{ii}}$. তাই $a_{ii} = p + iq (p, q \in \mathbb{R})$ হলে $p + iq = -p + iq$ হবে, তাই $p = 0$ হতে বাধ্য। **22.** $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$. **23.** $\left(A^{-1}\right)' = \left(A'\right)^{-1} = A^{-1}$, কারণ $A' = A$. **24.** আলবত! $\left(A^{-1}\right)^* = \left(A^*\right)^{-1} = A^{-1}$, যেহেতু $A^* = A$.

25. **(1)** $(A + A')' = A' + (A')' = A' + A = A + A'$.
(2) $(A - A')' = A' - (A')' = A' - A = -(A - A')$.
(3) $(A - A^*)^* = A^* - (A^*)^* = A^* - A = -(A - A^*)$. **(4)** $(AA^*)^* = (A^*)^*A^* = AA^*$.
(5) $(AA')' = (A')'A' = AA'$. **26.** ভুল, $-A$-ও symmetric হবে, কিন্তু skew symmetric নাও হতে পারে। যেমন $-\begin{bmatrix} 5 & 1 \\ 1 & 0 \end{bmatrix}$ মোটেই skew symmetric নয়। **27.**

$$\begin{aligned}(iA)^* &= (\overline{iA})' \\ &= -i\bar{A}' \\ &= -iA^* \\ &= -iA \quad [\because A^* = A.]\end{aligned}$$

Hence iA is skew-Hermitian.

28.

$$(B^*AB)^* = B^*A^*\left(B^*\right)^* = B^*A^*B.$$

Now if A is Hermitian, then $A^* = A$ and so $(B^*AB)^* = B^*AB$, and hence B^*AB is Hermitian.

Again, if A is skew Hermitian, then $A^* = -A$ and so $(B^*AB)^* = B^*(-A)B = -B^*AB$, and hence B^*AB is skew Hermitian.

29.

$$\begin{aligned}(AB + BA)^* &= (AB)^* + (BA)^* \\ &= B^*A^* + A^*B^* \\ &= BA + AB \quad [\because A^* = A, \quad B^* = B],\end{aligned}$$

as required.

$$\begin{aligned}(AB-BA)^* &= (AB)^*-(BA)^*\\ &= B^*A^*-A^*B^*\\ &= BA-AB \quad [\because A^*=A,\quad B^*=B]\\ &= -(AB-BA),\end{aligned}$$

as required.

30.

$$(I-A)(I+A)=I^2-A+A-A^2=I-A^2.$$

So

$$\begin{aligned}((I-A)(I+A))' &= (I-A^2)'\\ &= I'-(A^2)'\\ &= I-A'A'\end{aligned}$$

এবার দুটো case আলাদা করে ভাবতে হবে। এক, যদি A skew-symmetric হয়--

Now, if A is skew symmetric, then $A'=-A$.
So $A'A'=(-A)^2=A^2$.

আর, দুই, যদি A একটা symmetric matrix হয়--

Also, if A is symmetric, then $A'=A$.
So $A'A'=A^2$.

অতএব সিদ্ধান্তটা দুই ক্ষেত্রেই এক--

Hence in both the cases

$$((I-A)(I+A))'=I-A'A'=I-A^2=(I-A)(I+A).$$

So $(I-A)(I+A)$ is symmetric as required.

31.

Given $A^2-A-I=O$, or

$$A(A-I)=I \text{ and } (A-I)A=I.$$

Hence A^{-1} exists and equals $A-I$, as required.

32. Orthogonal-এর বেলায় $(A')'A'=AA'=I$ এবং $(-A)'(-A)=A'A=I$. আর unitary-র বেলায় $(A^*)^*A^*=AA^*=I$ এবং $(-A)^*(-A)=A^*A=I$. **33.** Orthogonal-এর বেলায়

$(AB)'(AB) = B'A'AB = B'IB = B'B = I$. একইভাবে unitary-র বেলাতেও হবে। **34.** (C). যদি matrix-টাকে A বলি, তবে $AA' = I$ হবে, ফলে $a_{ii}^2 = 1$. তাই n-খানা diagonal entry-র প্রত্যেকটাই 1 বা -1 যা খুশি হতে পারে। যেহেতু প্রতি ধাপেই 2-খানা করে সম্ভাবনা আছে (± 1), তাই উত্তরটা হল $\underbrace{2 \times \cdots \times 2}_{n} = 2^n$. **35.**

> We assume that A is a real matrix.
> Given
>
> $$\left((I+A)^{-1}(I-A)\right)' = -(I+A)^{-1}(I-A)$$
> or $$(I-A)'\left((I+A)^{-1}\right)' = -(I+A)^{-1}(I-A) \quad [\because (PQ)' = Q'P']$$
> or $$(I-A)'\left((I+A)'\right)^{-1} = -(I+A)^{-1}(I-A) \quad \left[\because (P')^{-1} = \left(P^{-1}\right)'\right]$$
> or $$(I-A')(I+A')^{-1} = -(I+A)^{-1}(I-A)$$
>
> Multiplying both sides by $(I+A')$ from right, and by $(I+A)$ from left,
>
> $$(I+A)(I-A') = -(I-A)(I+A')$$
> or $$I + A - A' - AA' = -I + A - A' + AA'$$
> or $$AA' = I.$$
>
> Since A is a square matrix, this also means $A'A = I$. Hence A is orthogonal, as required.

36.

> Since $(I-A)(I+A)^{-1}$ is orthogonal,
> so
>
> $$\left[(I-A)(I+A)^{-1}\right]'(I-A)(I+A)^{-1} = I$$
> or, $$\left[(I+A)^{-1}\right]'(I-A)'(I-A)(I+A)^{-1} = I$$
> $$[\because (PQ)' = P'Q']$$
> or, $$(I+A')^{-1}(I-A')(I-A)(I+A)^{-1} = I$$
> $$[\because (P^{-1})' = (P')^{-1}]$$
> or, $$(I-A')(I-A)(I+A)^{-1} = I + A' \quad \left[\text{multiplying by } I+A' \text{ from left}\right]$$
> or, $$(I-A')(I-A) = (I+A')(I+A) \quad \left[\text{multiplying by } I+A \text{ from right}\right]$$
> or, $$I - A' - A + A'A = I + A' + A + A'A$$
> or, $$-A' - A = A' + A$$
> or, $$A' = -A,$$
>
> ie, A is skew-symmetric, as required.

37.

First we notice that

$$(I+A)^{-1}(I-A) = (I-A)(I+A)^{-1}.$$

⟦ Because:

By direct multiplication

$$(I-A)(I+A) = (I+A)(I-A).$$

So multiplying by $(I+A)^{-1}$ from left and right, the result follows.

⟧

Also $(I+A') = (I+A)'$ is nonsingular, since $(I+A)$ is nonsingular.

Now

$$\begin{aligned}
(I+A)^{-1}(I-A) &= (AA'+A)^{-1}(I-A) \quad [\because AA'=I]\\
&= [A(A'+I)]^{-1}(I-A)\\
&= (A'+I)^{-1}A^{-1}(I-A)\\
&= (A'+I)^{-1}(A^{-1}-I)\\
&= (A'+I)^{-1}(A'-I) \quad [\because A = A^{-1}]\\
&= [(A+I)']^{-1}(A-I)'\\
&= [(A+I)^{-1}]'(A-I)'\\
&= [(A-I)(A+I)^{-1}]'\\
&= -[(I-A)(A+I)^{-1}]'\\
&= -[(A+I)^{-1}(I-A)]'.
\end{aligned}$$

Hence the result.

38. এখানে induction on n লাগাতে হবে। যদি $n=1$ তবে তো বলাই বাহুল্য, কারণ $A^1 = A$. যদি $A^n = A$ ধরে নাও, তবে $A^{n+1} = A^{n-1}A^2 = A^{n-1}A = A^n = A$. **39.** যোগ আর গুণের নামতা হল

$\oplus_2$	0	1
0	0	1
1	1	0

এবং

$\otimes_2$	0	1
0	0	0
1	0	1

. এখানে $1 \odiv_2 (1 \oplus_2 1 \oplus_2 1) = 1 \odiv_2 1 = 1$. **40.** $A^{-1} = \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}$.

41. $AB = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$.

Chapter II
Vector space

DAY 7
Vector space

আমরা গত অধ্যায়ে ভূতের রাজার বরের কথা বলেছিলাম মনে আছে? শেখার সুবিধার জন্য linear algebra বিষয়টাকে আধুনিককালে সম্পূর্ণ ঢেলে সাজানো হয়েছে। গত অধ্যায়ে আমরা সেই আধুনিক পথে শুরু করিনি, বরং হায়ার সেকণ্ডারীতে যা যা শিখেছিলাম, সেইগুলোকেই একটু ঝালিয়ে নিচ্ছিলাম। সেই অধ্যায়ের শেষে যে field-এর কথা আলোচনা করেছিলাম, সেখান থেকেই আধুনিক পথের শুরু। সেই পথের পরবর্তী ধাপ হল vector space নামে একটা জিনিস।

Vector শব্দটা শুনলেই তোমার আবার হায়ার সেকণ্ডারীর স্মৃতি জেগে উঠতে পারে, যেখানে আমরা vector বলতে ছোটো ছোটো তীর চিহ্ন আঁকতাম, যেমন ধরো বৃষ্টির ফোঁটা পড়ছে, তার velocity বোঝানোর জন্য একটা তীর চিহ্ন (Fig 1), সেই বৃষ্টির মধ্যে দিয়ে একজন হেঁটে যাচ্ছে, তার velocity-র জন্য আরেকটা তীর চিহ্ন, এইরকম। কিংবা টেবিলের উপর একটা বাক্স আছে, বাক্সের ওজন বোঝাতাম নীচের দিকে একটা তীর চিহ্ন টেনে (Fig 2)। আবার টেবিলটা যে বাক্সটাকে পড়তে দিচ্ছে না, তার জন্য যে force-টা, সেটাকে দেখাতাম উপর দিকে একটা তীর চিহ্ন এঁকে। ছবিতে এই তীর চিহ্ন দুটোর নাম দিয়েছি $\vec{W}$ আর $\vec{R}$. উপরের ওই ছোট্টো তীর চিহ্নটাই বলে দিচ্ছে যে ওটা একটা vector.

আমরা এখানে যে vector-এর কথা শিখব সেটাও একইরকম জিনিস, খালি একটু বেশী general. হায়ার সেকণ্ডারীতে শেখা vector-গুলো হবে এদের একটা special case. তীর চিহ্ন আমরা এখনও আঁকব, তবে আরও এমন সব vector-এর কথা জানতে পারব, যাদের সঙ্গে তীর চিহ্নর আপাতদৃষ্টিতে কোনোই সম্পর্ক নেই। কিন্তু সেই general আলোচনায় যাওয়ার আগে আমরা হায়ার সেকণ্ডারীর স্মৃতিটা আরেকটু রোমন্থন করে নিই।

টেবিলের উপর বাক্সের ওজন দেখিয়েছিলাম নীচের দিকে তীর চিহ্ন দিয়ে। এবার মনে করো ওর বদলে একটা দ্বিগুণ ভারী বাক্স বসালাম। তবে তীর চিহ্নটা কীভাবে বদলাবে? উত্তর হল--vector-টার অভিমুখ একই থাকবে, কিন্তু লম্বায় দ্বিগুণ হয়ে যাবে (Fig 3). সুতরাং একটা vector-কে 2 দিয়ে গুণ করা মানে তীরটার অভিমুখ না বদলে লম্বায় টেনে দ্বিগুণ করে দেওয়া। খালি 2 বলেই নয়, যাই সংখ্যা দিয়ে গুণ কর ব্যাপারটা একইরকম। যেমন আগের বাক্সটার কথাই ধরো। যেহেতু বাক্সটা টেবিলের উপর স্থির হয়ে আছে, তার মানে টেবিলটা বাক্সটাকে ঠেলে রাখছে ঠিক সমান force দিয়ে, তবে উল্টো দিকে। সুতরাং ওজনটাকে যদি $\vec{F}$ বলি, তবে এই force-টাকে $-\vec{F}$ বলা চলে (Fig 4)। মানে কোনো vector-কে মাইনাস করলে সেটা বেমালুম উল্টে যায়। বুঝতেই পারছ যে -2 দিয়ে গুণ করলে কী হত--উল্টে তো যেতই, তার উপরে আবার লম্বাতেও

Fig 1



Fig 2

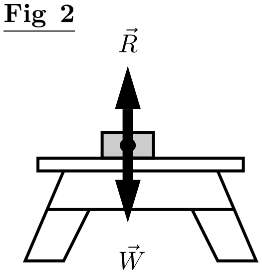


Fig 3

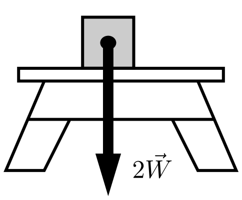
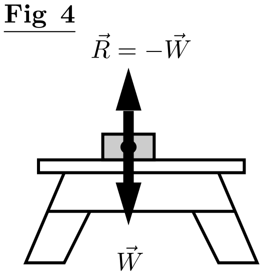


Fig 4



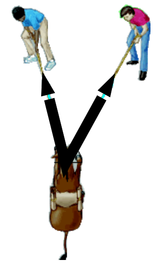


Fig 5

দ্বিগুণ হয়ে যেত। মানে -2 দিয়ে গুণ করাটাকে দুই ধাপে ভেঙে ভাবা যায়--প্রথমে মাইনাস করে উল্টে দেওয়া, তারপর 2 দিয়ে গুণ করে লম্বা করে দেওয়া--

$$-2\vec{F} = (2\times -1)\vec{F} = 2(-\vec{F}).$$

এসবই তুমি হায়ার সেকণ্ডারী থেকেই জানো। আরও একটা জিনিস জানো--সেটা হল কী করে কিছু vector-কে যোগ করে resultant (লব্ধি) বার করতে হয়। এর জন্য আমরা ব্যবহার করতাম parallelogram law. যেমন ধরো একটা ঘোড়াকে দুজনে দু দিকে টানছে (Fig 5)। ঘোড়াটা তাহলে কোন দিকে যাবে? এর জন্য আমরা একটা parallelogram আঁকি, এবং তার diagonal-টা নিই (Fig 6)। এইভাবে ছবি আঁকাটা ভালোই, তবে ব্যাপারটাকে সংখ্যা দিয়ে প্রকাশ করে নিলে সুবিধা হবে। এর জন্য ধরো তীর চিহ্ন দুটো একটা গ্রাফকাগজে আঁকা আছে (Fig 7). গোড়াদুটো রয়েছে origin-এ। তাহলে ডগাদুটোর অবস্থান জানলেই তীর চিহ্নদুটোর অবস্থান জানা হবে। ধরো ডগাদুটো আছে (a,b) আর (c,d)-তে। তাহলে একটু চিন্তা করলেই বুঝবে যে resultant-এর ডগাটা থাকবে $(a+c,b+d)$-তে। এই জিনিসটা কিন্তু খুবই গুরুত্বপূর্ণ। কারণ এর ফলে দুটো vector-এর যোগ পরিণত হল কিছু সাধারণ সংখ্যার যোগে! এতে একটা সুবিধা হল এই যে, সাধারণ যোগের যা যা বৈশিষ্ট্য আমরা জানি সেগুলো vector-এর যোগের বেলাতেও লাগানো যাবে। যেমন ধরো আমরা জানি যে x,y দুটো সংখ্যা হলে $2(x+y)=2x+2y$ হয়। একই যুক্তিতে আমরা বলতে পারব যে $2(\vec{F}_1+\vec{F}_2)=2\vec{F}_1+2\vec{F}_2$ হবে। সুতরাং লোক দুটো যদি দ্বিগুণ জোরে টানে, তবে ঘোড়াটা একই দিকে দ্বিগুণ জোরে ছিটকে পড়বে (Fig 8)।

এই গেল হায়ার সেকণ্ডারীর স্মৃতিরোমন্থন। আমরা এখানে তীর চিহ্নদের নিয়ে অংক করলাম--তীর চিহ্নদের সংখ্যা দিয়ে গুণ করা, দুটো তীর চিহ্নকে যোগ বিয়োগ করা। যদিও "তীর চিহ্ন" জিনিসটা ঠিক একটা সংখ্যা নয়, কিন্তু দেখা গেল যে ওদের যোগ-বিয়োগ ইত্যাদির ধর্মগুলো অনেকটা সাধারণ সংখ্যার যোগ-বিয়োগের মতই। এই যে ব্যাপারটা-- "সংখ্যা নয়, অথচ সংখ্যার মতই যোগ-বিয়োগ ইত্যাদি করা যায়" এই ধারণাটা তীর চিহ্ন ছাড়া অন্যত্রও দেখা যায়।

Example 1: যেমন ধরো যাবতীয় 2×3 matrix-রা। (অবশ্য 2×3 ছাড়া অন্য যা খুশি সাইজও নিতে পারো।) এরকম যেকোনো matrix-কে তুমি কোনো সংখ্যা দিয়ে গুণ করতে পারো। একই সাইজের দুটো matrix-কে যোগ বিয়োগও করতে পারো। এখানেও যোগ-বিয়োগ ইত্যাদির নিয়ম সাধারণ সংখ্যার মতই। যেমন $2(A+B)=2A+2B$. লক্ষ কর এখানে আমরা কিন্তু দুটো matrix-কে গুণ করার প্রসঙ্গে যাচ্ছি না। বস্তুতঃ দুটো 2×3 সাইজের matrix-কে গুণ করা যায়ই না। ■

যদিও matrix-দের সঙ্গে তীর চিহ্নের কোনোই মিল নেই তাও কেমন দ্যাখো অংক কষার দিক থেকে ওদের মধ্যে আশ্চর্য

Fig 6

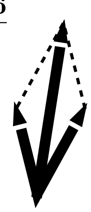


Fig 7

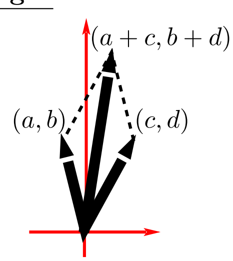
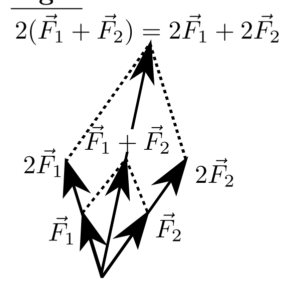


Fig 8



মিল! এবার আরেকটা পরিচিত জিনিসের কথা বলব যেটা matrix বা তীর চিহ্ন কোনোটার মতই না, অথচ যোগ-বিয়োগ ইত্যাদির ব্যাপারটা একইরকম।

Example 2: ধরো একটা function আছে $f(x)$. যেমন $\sin x$ বা $\cos x$ বা $\log x$, এই রকম কিছু একটা। তুমি কোনো function-কে একটা সংখ্যা দিয়ে গুণ করতে পারো, যেমন $\sin x$-কে 2 দিয়ে গুণ করলে হয় $2\sin x$. আবার দুটো function দেওয়া থাকলে তাদের যোগ-বিয়োগও করা যায়, যেমন $\sin x+\log x$ বা $\cos x-e^x$. এখানেও যোগ-বিয়োগের সাধারণ নিয়মগুলো খাটে। যেমন $\sin x+\cos x=\cos x+\sin x$ হয়, বা

$$-1\times 2\times \sin x=-1\times(2\sin x)$$

হয়, ইত্যাদি। ■

এই অব্ধি পড়ে তোমার হয় তো মনে হচ্ছে যে, খামোখা আমরা মামূলী কথাগুলোকে সাতকাহন করে বলছি কেন? তার কারণ হল দেখা গেছে linear algebra যেখানে যেখানে কাজে লাগে, সব ক্ষেত্রেই এই যোগ-বিয়োগের ব্যাপারটা থাকে। সেই কারণে এই রকম যাবতীয় ক্ষেত্রের একটা নাম দেওয়া হয়েছে--**vector space.** এবার এই জিনিসটার সংজ্ঞাই আমরা শিখব। সংজ্ঞাটার প্রথম উপাদান হল একটা set. কোন্ ধরণের জিনিসকে যোগ বিয়োগ করব তাদের set এটা। যেমন উপরের উদাহরণগুলোতে যাবতীয় তীর চিহ্নের set, বা যাবতীয় 2×3 সাইজের matrix-এর set, বা যাবতীয় function-এর set, এইরকম। এই set-এর element-দের আমরা বলব একেকটা **vector.** এদের নিয়ে দুই ধরণের কাজ করা যায়, কোনো সংখ্যা দিয়ে গুণ করা, আর দুটো vector-এর যোগ-বিয়োগ করা। কোনো সংখ্যা দিয়ে গুণ করার উদাহরণ যেমন একটা তীরকে 2 দিয়ে গুণ করলে তীরটা লম্বায় দ্বিগুণ হয়, বা $\sin x$-কে 2 দিয়ে গুণ করলে $2\sin x$ হয়, এইরকম। এখানে অবশ্যই প্রশ্ন ওঠে যে সংখ্যা বলতে আমরা কী বোঝাচ্ছি? এই একই প্রশ্নের সম্মুখীন আমরা প্রথম অধ্যায়েও হয়েছিলাম--একটা matrix-এর entry হিসেবে কী ধরণের সংখ্যা নেওয়া চলে? সেখানে যে উত্তরটা দিয়েছিলাম, এখানেও সেই উত্তরেই কাজ চলবে--আমরা গোড়াতেই কোনো একটা field ঠিক করে নিয়ে, "সংখ্যা" বলতে তার element-দেরই বুঝব। এই "সংখ্যা"-দের পোশাকি নাম যে **scalar** সে তো গত অধ্যায়েই বলেছি। কী field নিচ্ছি সেটা অংকের শুরুতেই লিখে দিতে হয়, যেমন যদি লেখা দ্যাখো--"Let V be a vector space over the field $\mathbb{Q}$" অমনি বুঝবে যে এখানে rational number-দের নিয়ে কাজ হবে। যদি field-টা হয় $\mathbb{R}$, তবে লোকে অনেক সময়ে বলে "V is a real vector space," মানে "V is a vector space over the field $\mathbb{R}$." অবশ্য অনেক সময়েই আমরা এমন আলোচনা করব যা যেকোনো field-এর ক্ষেত্রেই সমানভাবে প্রযোজ্য, সেক্ষেত্রে $\mathbb{Q},\mathbb{R}$ বা $\mathbb{C}$ ইত্যাদি আলাদা নাম না ব্যবহার করে, খালি F লিখে ছেড়ে দেব এইভাবে--"V is a vector space over some field F." অর্থাৎ তুমি F-এর জায়গায় $\mathbb{Q},\mathbb{R},\mathbb{C}$ বা কোনো finite field ইত্যাদি যাই কল্পনা করো না কেন, আমাদের আলোচনা একই থাকবে। তবে প্রথম প্রথম linear algebra শেখার সময়ে field-এর বৈচিত্র্য নিয়ে মাথা না ঘামানোই ভালো, খালি $\mathbb{R}$ (বা বড়জোর $\mathbb{C}$) ধরে এগোলেই সুবিধা হয়। সেক্ষেত্রে scalar-রা আসলে আমাদের পরিচিত সংখ্যারাই হবে।

কোনো scalar দিয়ে একটা vector-কে গুণ করা মানে হল $a\cdot\vec{v}$ বার করা, যেখানে $a\in F$ হল scalar আর $\vec{v}\in V$ হল vector. এই গুণফলটা নিজে আবার একটা vector হবে। সুতরাং বুঝতেই পারছ যে এই গুণটা হল একটা function যেটা input হিসেবে একটা scalar আর একটা vector নেয়, এবং তাদের গুণ করে output একটা vector তৈরী করে। তার মানে এটা হল $F\times V$ থেকে V-তে একটা function. আমরা $a\cdot\vec{v}$ লেখার সময়ে সব সময়ে প্রথমে scalar-টা লিখি এবং পরে vector-টা লিখি। তাড়াতাড়ি লেখার সময়ে প্রায়ই ওই মাঝের ডট্‌টা বাদ দিয়ে খালি $a\vec{v}$ লিখে ছেড়ে দিই। এই গুণটাকে বলা হয় **scalar multiplication**. এই গেল গুণের কথা। এবার দুটো vector-কে যোগ করা নিয়ে আলোচনা করি।

ধরো $\vec{v}$ আর $\vec{w}$ দুটো vector আছে, তবে তাদের যোগফল হল $\vec{v}+\vec{w}$. তার মানে $+$ হল একটা function, যেটা দুটো vector-কে মিলিয়ে একটা vector-এ পরিণত করে। অর্থাৎ $+$ হল $V\times V$ থেকে V-তে একটা function. এরপর আসে বিয়োগের কথা। চাইলে আমরা তার জন্য আরেকটা function বানাতে পারি। কিন্তু সাধারণতঃ সেটা না করে বিয়োগকেও যোগ দিয়েই প্রকাশ করা হয়, যেমন সাধারণ সংখ্যার ক্ষেত্রেও করা যায়, $3-2=3+(-2)$. এখানে -2 হল সেই সংখ্যা যাকে 2-এর সাথে যোগ করলে 0 হয়। Vector space-এর বেলাতেও আমরা একইভাবে এগোই--$\vec{v}\in V$-এর জন্য আমরা $-\vec{v}$ নিই সেই vector-কে যাতে $\vec{v}+(-\vec{v})=\vec{0}$ হয়। বুঝতেই পারছ এর জন্য একটা zero vector চাই, যাকে $\vec{0}$ লিখেছি। এবার যদি $\vec{u}-\vec{v}$ লিখতে হয়, তবে সেটাকে আমরা $\vec{u}+(-\vec{v})$ হিসেবে লিখতে পারি। সুতরাং আমাদের হাতে এবার যোগ বিয়োগ আর scalar multiplication লেখার হাতিয়ার এসে গেছে, এবার খালি বলে দিতে হবে যে এই যোগ-বিয়োগ ইত্যাদিরা

সাধারণ সংখ্যার যোগ বিয়োগের মতই আচরণ করবে। এই কাজটা করলেই আমরা vector space-এর সংজ্ঞা পেয়ে যাব। ব্যাপারটা অনেকটা field-এর সংজ্ঞা লেখার মতই, ধর্মগুলোও অনেকটা একইরকম, যদিও কিছু গুরুত্বপূর্ণ পার্থক্য আছে।

DEFINITION: Vector space

Let F be a field. Then a vector space over F means a set V together with two operations

- $+ : V \times V \to V$ called vector addition,
- $\cdot : F \times V \to V$ called scalar multiplication,

such that the following conditions hold:

1. Closure under $+$: $\forall \vec{u}, \vec{v} \in V \quad \vec{u} + \vec{v} \in V$
2. Associativity of $+$: $\forall \vec{u}, \vec{v}, \vec{w} \in V \quad (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$.
3. Commutativity of $+$: $\forall \vec{u}, \vec{v} \in V \quad \vec{u} + \vec{v} = \vec{v} + \vec{u}$
4. Additive identity: $\exists \vec{0} \in V \quad \forall \vec{v} \in V \quad \vec{v} + \vec{0} = \vec{v}$
5. Additive inverse: $\forall \vec{v} \in V \quad \exists \vec{w} \in V \quad \vec{v} + \vec{w} = \vec{0}$. We denote this $\vec{w}$ by $-\vec{v}$.
6. Closure under $\cdot$: $\forall c \in F \quad \forall \vec{v} \in V \quad c \cdot \vec{v} \in V$.
7. $\forall a, b \in F \quad \forall \vec{v} \in V \quad (ab) \cdot \vec{v} = a \cdot (b \cdot \vec{v})$.
8. $\forall \vec{v} \in V \quad 1 \cdot \vec{v} = \vec{v}$, where 1 is the multiplicative identity of F.
9. Distributivity:
   - $\forall a, b \in F \quad \forall \vec{v} \in V \quad (a + b) \cdot \vec{v} = (a \cdot \vec{v}) + (b \cdot \vec{v})$.
   - $\forall a \in F \quad \forall \vec{u}, \vec{v} \in V \quad a \cdot (\vec{u} + \vec{v}) = (a \cdot \vec{u}) + (a \cdot \vec{v})$.

সংজ্ঞাটায় একরাশ শর্ত আছে, আর কিছু খটমট শব্দ আছে। সেগুলো হজম করে নেওয়া যাক। শর্তগুলো সবই সাধারণ সংখ্যার ক্ষেত্রে যোগ আর গুণের যে সব ধর্মের সঙ্গে আমরা ছোটোবেলা থেকেই পরিচিত, তাদেরই মত। প্রথম পাঁচটা শর্ত হল $+$ নিয়ে। পরবর্তী তিনটে গুণ নিয়ে, এবং সবশেষের শর্তটা যোগ আর গুণের সম্পর্ক নিয়ে। এই সংজ্ঞার সঙ্গে field-এর সংজ্ঞার তুলনা করা যাক।

- দুটো scalar-এর যোগ আর দুটো vector-এর যোগ, এই দুইরকম যোগেরই ধর্মগুলো একই। দুই ক্ষেত্রেই "শূন্য"-র ধারণা আছে। দুটো শূন্য কিন্তু আলাদা, scalar শূন্যটাকে লিখি 0, আর vector শূন্যটাকে লিখি $\vec{0}$.
- একইভাবে, যোগের উল্টো যে বিয়োগ, সেটাও field এবং vector space দুয়ের ক্ষেত্রেই আছে। কোনো কিছুর আগে মাইনাস চিহ্নটা বসানোকে দুইক্ষেত্রেই বলে **additive inverse** বা **negative** বার করা।
- গুণের বেলায় field আর vector field অনেকটাই আলাদা। Field-এর গুণটা হল দুটো scalar-এর মধ্যে। কিন্তু vector space-এর গুণটা দুটো vector-এর মধ্যে নয়, একটা scalar আর একটা vector-এর মধ্যে। সেই কারণে field-এর মধ্যে যেমন একটা 1 ছিল, vector space-এর বেলায় সেরকম কিছু নেই।

7.1 নানা ধরণের vector space

কিছু vector space-এর উদাহরণ দেখব এইবার। প্রথমেই আসবে সেই তীর চিহ্নদের উদাহরণ। আমরা আগেই দেখেছি তীর চিহ্নগুলোকে গ্রাফ কাগজে আঁকলে ওদেরকে (x, y) হিসেবে ভাবা যায়। এবং সেক্ষেত্রে scalar multiplication-টা হয় $a(x, y) = (ax, ay)$ আর যোগের নিয়মটা হয় $(x, y) + (u, v) = (x + u, y + v)$. এখানে আমাদের field-টা হল $\mathbb{R}$, মানে a, x, y, u, v এরা সবাই real number. যেহেতু প্রত্যেকটা vector-কে দেখতে (x, y)-এর মত, যার মধ্যে 2-খানা real number আছে, তাই এই vector space-টাকে লেখা হয় $\mathbb{R}^2$. তীরগুলো যদি খাতার পাতায় আঁকা হয়, তবে খালি x আর y-তেই কাজ চলে যাবে। কিন্তু যদি 3-dimensional (ত্রিমাত্রিক) হয়, তবে লাগবে (x, y, z). সেক্ষেত্রেও scalar multiplication আর vector addition-এর নিয়ম একইরকম থাকবে--

$$\begin{aligned} a(x, y, z) &= (ax, ay, az), \\ (x_1, y_1, z_1) + (x_2, y_2, z_2) &= (x_1 + x_2, y_1 + y_2, z_1 + z_2). \end{aligned}$$

এখানেও field-টা নেব $\mathbb{R}$. বুঝতেই পারছ যে এই vector space-টার নাম $\mathbb{R}^3$. একইভাবে তুমি $\mathbb{R}^4, \mathbb{R}^5, \ldots$ ইত্যাদিও বানাতে পারো। এদের general চেহারাটা হল $\mathbb{R}^n$. এখানে প্রত্যেকটা vector দেখতে হয় এইরকম--

$$(x_1, \ldots, x_n),$$

যেখানে প্রত্যেকটা $x_i \in \mathbb{R}$. এদের scalar multiplication এবং vector addition করা হয় এইভাবে--

$$\begin{aligned} a(x_1, \ldots, x_n) &= (ax_1, \ldots, ax_n) \\ (x_1, \ldots, x_n) + (y_1, \ldots, y_n) &= (x_1 + y_1, \ldots, x_n + y_n). \end{aligned}$$

$\mathbb{R}^2$ বা $\mathbb{R}^3$-র vector-দের তীর চিহ্ন বলে ভাবা যায়, কিন্তু অনেক ছাত্রছাত্রীর মনে প্রশ্ন ওঠে যে $n > 3$ হলে $\mathbb{R}^n$ নিয়ে আমরা কেন মাথা ঘামাব? বাস্তবে কি ওদের কোনো প্রয়োগ আছে, নাকি ওরা নিছক অংকের কল্পনাবিলাস? উত্তর হল--চিন্তা নেই, বাস্তবে ওদের বিস্তর প্রয়োগ আছে। যেমন ধরো নীচের উদাহরণটা--

Example 3: একটা ভিটামিন ট্যাবলেটে বিভিন্ন ভিটামিনের পরিমাণ এইরকম--

ভিটামিন	A	B	C	D	E	K
পরিমাণ	0.1	0.01	0.5	0	0	0

আরেক ধরণের ভিটামিন ট্যাবলেটের উপাদানগুলো এইরকম--

ভিটামিন	A	B	C	D	E	K
পরিমাণ	0	0.2	0.1	0.1	0.3	0.01

লক্ষ কর যে দুটো তালিকাই বস্তুতঃ 6-টা করে সংখ্যা দিয়ে তৈরী। এইরকম ক্ষেত্রে ওদেরকে আমরা $\mathbb{R}^6$-এর দুটো vector বলে ভাবতে পারি--

$$\vec{v}_1 = (0.1, 0.01, 0.5, 0, 0, 0), \quad \vec{v}_2 = (0, 0.2, 0.1, 0.1, 0.3, 0.01).$$

এবার যদি কেউ প্রথম ট্যাবলেটের আধখানা আর দ্বিতীয় ট্যাবলেটের আধখানা খায় তবে সে কী পরিমাণ ভিটামিন পাবে? উত্তর হল--

$$\tfrac{1}{2}\vec{v}_1 + \tfrac{1}{2}\vec{v}_2 = (0.05, 0.105, 0.3, 0.05, 0.15, 0.005).$$

দেখলে, এখানে কীভাবে scalar multiplication এবং vector addition ব্যবহার হল? ■

আরেকটা উদাহরণ দিই, এবার physics থেকে--

Example 4: একটা পেণ্ডুলাম দুলছে, আমরা ক্যামেরা দিয়ে কোনো এক মুহূর্তে তার ছবি তুলেছি। ছবিটা দেখতে হয়েছে

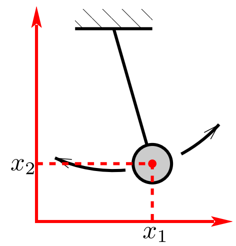


Fig 9

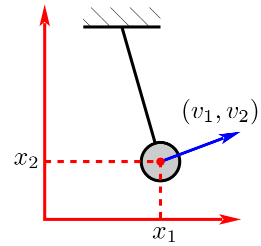


Fig 10

Fig 9-এর মত। এই ছবি থেকে আমরা সেই মুহূর্তে পেণ্ডুলামের অবস্থানটা চমৎকার বুঝতে পারছি। আমরা যদি মনে মনে গ্রাফ কাগজের মত দুটো axis কল্পনা করে নিই, তবে পেণ্ডুলামের অবস্থানটাকে (x_1, x_2) হিসেবে লেখা যায়। কিন্তু খালি (x_1, x_2) থেকেই কি বলতে পারবে পেণ্ডুলামটা সেই মুহূর্তে কীভাবে দুলছিল? না, এই একটামাত্র স্থিরচিত্র থেকে সেটা বলা অসম্ভব। যদি পেণ্ডুলামটার গতির পুরো বর্ণনাটা চাও, তবে খালি অবস্থান (x_1, x_2) জানলেই চলবে না, তার বর্তমান velocity-টাও জানতে হবে, (v_1, v_2), যেমন দেখিয়েছি Fig 10-এ। তার মানে পেণ্ডুলামের গতির সম্পূর্ণ বর্ণনাটা চারটে সংখ্যা দিয়ে তৈরী--(x_1, x_2, v_1, v_2). তাই একজন physicist-এর চোখে পেণ্ডুলামের গতিটা রয়েছে $\mathbb{R}^4$-এর মধ্যে। ■

Exercise 1: আচ্ছা বল তো, যদি $n = 1$ নিই, তবে $\mathbb{R}^n$ আসলে কী? ■

আমরা এতক্ষণ দেখলাম যে $\mathbb{R}^n$ হল $\mathbb{R}$-এর উপর একটা vector space. এখানে আমরা $\mathbb{R}$-এর জায়গায় $\mathbb{C}$ নিয়েও কাজ করতে পারতাম, তাহলে পেতাম $\mathbb{C}^n$ যেটা $\mathbb{C}$-এর উপরে একটা vector space. উদাহরণস্বরূপ $\mathbb{C}^2$-এর একজন সদস্য হবে এইরকম--

$$(2 + 3i, 5 - 6i).$$

শুধু $\mathbb{R}$ বা $\mathbb{C}$-ই কেন, চাইলে তুমি যেকোনো field নিয়েই একই কাজ করতে পারো। যেমন যদি field-টার নাম দিই F, তবে তার উপরে একটা vector space হবে F^n. এখানে প্রতিটা vector হল $(f_1, ..., f_n)$-এর মত যেখানে f_i-রা সবাই F-এর element.

আমরা অনেক অংকের শুরুতেই বলি--Let $\mathbb{R}^n$ be a vector space." এখানে আমরা field-টা কী নিচ্ছি সেটা আলাদা করে না বললেও বুঝে নিতে হবে যে $\mathbb{R}$-এর উপরেই কাজ হচ্ছে। একইভাবে যদি vector space-টা হয় $\mathbb{C}^n$, আর field-টার উল্লেখ না করা থাকে, তবে বুঝে নেবে যে এখানে field-টা $\mathbb{C}$. অর্থাৎ F^n নিয়ে কাজ হলে field-টা লোকে সাধারণতঃ F-ই নেয়। এই প্রসঙ্গে নীচের সামান্য ধাঁধানো প্রশ্নটা মাথায় এল--

Exercise 2: নীচের দুটো বাক্যের মধ্যে একটাই খালি সঠিক, কোনটা?

1. $\mathbb{R}^n$ is a vector space over $\mathbb{C}$.
2. $\mathbb{C}^n$ is a vector space over $\mathbb{R}$.

■

Example 5: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). Which one is not a vector space?

(A) $\mathbb{C}$ over $\mathbb{C}$ (B) $\mathbb{C}$ over $\mathbb{R}$ (C) $\mathbb{R}$ over $\mathbb{C}$ (D) $\mathbb{R}$ over $\mathbb{Q}$

(2021.CC6.1h)
SOLUTION:chap 2/ day 7 ■

এবার একটা নতুন ধরণের vector space-এর উদাহরণ দেখব। তার আগে notation নিয়ে একটা কথা বলে নিলে সুবিধা হবে। আমরা যখন কোনো function বোঝাতে চাই, তখন সাধারণতঃ $f(x)$ লিখি, যেন $f(x)$ হল function-টার নাম। কিন্তু একটু ভাবলেই বুঝবে যে function-টার নাম খালি f, এখানে $f(x)$ হল x-এ f-র value. যদি আরেকটা function নিই g, তবে x-এ g-এর value হবে $g(x)$. এবার ধরো f আর g-কে যোগ করে দিলাম। তবে একটা নতুন function পাব, তাকে আমরা লিখব $f+g$. এবার যদি x-এ আমরা $f+g$-এর value বার করতে চাই (মানে $(f+g)(x)$ বার করতে চাই), তার উত্তর হল--প্রথমে f আর g-এর value আলাদা করে বার করে যোগ করে দিতে হবে, মানে

$$(f+g)(x) = f(x) + g(x).$$

এই $(f+g)(x)$ দেখে অনেক ছাত্র একটু খাবি খায়। এর মানে আসলে $f(x)+g(x)$ ছাড়া আর কিছুই নয়। ঠিক যেমন $\sin x + \cos x$ না লিখে $(\sin + \cos)(x)$ লেখা। এর একটা গালভরা নামও আছে--একে বলে **pointwise addition,** অর্থাৎ কিনা f আর g-কে প্রত্যেকটা point-এ (মানে x-এর প্রতিটা value-তে) আলাদা করে যোগ করা হয়েছে।

Example 6: Let $C[a,b]$ be the set of all real valued continuous functions on $[a,b]$. Let $f(x)+g(x) = (f+g)(x)$ and $\alpha f(x) = (\alpha f)(x)$ for $f,g \in C[a,b]$ and $\alpha \in \mathbb{R}$. Show that $C[a,b]$ is a vector space over $\mathbb{R}$. **(2008.9a)**
SOLUTION:

> We know that sum of continuous functions is again continuous. Also scalar multiples of continuous functions are continuous. So addition and scalar multiplication are well-defined.

এখানে আমরা "We know that..." বলেছি কারণ continuous function-দের এইসব ধর্ম আমরা হায়ার সেকণ্ডারীতে এবং real analysis শেখার সময়েই শিখি। "Multiple" শব্দটার অর্থ হল "গুণিতক", যেমন 10 হল 5-এর একটা multiple. এখানে অবশ্য function-এর multiple-এর কথা বলেছে। যেমন, $\sin x$-এর একটা scalar multiple হল $2\sin x$.

> Associativity of $+$:

দেখাতে হবে যে, যদি $f,g,h \in C[a,b]$ হয় তবে $(f+g)+h = f+(g+h)$ হবে, মানে যাই $x \in [a,b]$ নাও না কেন, $((f+g)+h)(x) = (f+(g+h))(x)$ হবে।

> For $f,g,h \in C[a,b]$ and $x \in [a,b]$ we have
>
> $$\begin{aligned} ((f+g)+h)(x) &= (f+g)(x)+h(x) \\ &= (f(x)+g(x))+h(x) \\ &= (f(x)+(g(x)+h(x)) \quad [\because + \text{ is also in } \mathbb{R}] \\ &= f(x)+(g+h)(x) \\ &= (f+(g+h))(x). \end{aligned}$$
>
> Commutativity of $+$: For $f,g \in C[a,b]$ and $x \in [a,b]$ we have
>
> $$(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x),$$

Since $+$ is commutative in $\mathbb{R}$.
Additive identity: The function $z(x) \equiv 0 \in C[a, b]$ is the additive identity.
Additive inverse: For $f \in C[a, b]$, the function $-f \in C[a, b]$ defined by

$$(-f)(x) = -f(x)$$

is its additive inverse.

এবার scalar multiplication-এর ধর্মগুলো দেখাব। Closure তো দেখানো হয়েই গেছে।

Since multiplication in $\mathbb{R}$ is associative, for any $f \in C[a, b]$ and any $a, b \in \mathbb{R}$

$$((ab)f)(x) = (ab)f(x) = a(bf(x)) = a(bf)(x) = (a(bf))(x).$$

1 দিয়ে গুণ করলে কিছু বদলায় না--

Since 1 is the multiplicative identity in $\mathbb{R}$,

$$(1f)(x) = 1f(x) = f(x).$$

পড়ে থাকে খালি দুই ধরণের distributivity দেখানো--

Distributivity: Since multiplication distributes over addition in $\mathbb{R}$, so for any $f, g \in C[a, b], a, b \in \mathbb{R}$ we have

$$\begin{aligned} (a(f+g))(x) &= a(f+g)(x) = a(f(x)+g(x)) \\ &= af(x) + ag(x) = (af+ag)(x) \\ ((a+b)f)(x) &= (a+b)f(x) = af(x) + bf(x) \\ &= (af+bf)(x). \end{aligned}$$

So $C[a, b]$ is a vector space over $\mathbb{R}$ under the given operations.

■

এই অংকটা ভালো করে দেখে নাও। এখানে continuous শর্তটা খালি দু জায়গায় লেগেছে-- এক, যোগ আর গুণকে well-defined দেখাতে, আর দুই, additive identity বানানোর সময়ে। এই কথাটা মাথায় রেখে নীচের অংকটা কর।

Exercise 3: Let P stand for the set of all functions from $[0, 1]$ into $\mathbb{R}$, which are differentiable over $[0, 1]$. Show that P is a vector space over $\mathbb{R}$ if addition of functions in P and multiplication of functions in P by the elements of $\mathbb{R}$ are defined pointwise on $[0, 1]$.[3] **(2009.9a)**
HINT: ঠিক আগের অংকটার মতই এগোও। খালি closure-এর জন্য এই কথাটা ব্যবহার কর যে, দুটো differentiable function-এর যোগফলও differentiable হয় (যেমন $\sin x + \cos x$ একটা differentiable function), এবং কোনো differentiable function-কে কোনো সংখ্যা দিয়ে গুণ করলে গুণফলটাও differentiable হয় (যেমন $2 \sin x$ একটা differentiable function). আর additive identity-র জন্য লক্ষ্য কর যে, $f(x) \equiv 0$ একটা differentiable function. ■

এবার vector space-এর একটা সহজ ধর্ম শিখব নীচের অংকটায়।

Example 7: In a real vector space V prove that

$$(-c)\alpha = -(c\alpha) = c(-\alpha),$$

where $c \in \mathbb{R}$ and $\alpha \in V$.[2] **(2006.13cii)**
SOLUTION: জিনিসটা দেখতে বেশ সহজ, ঠিক যেন "মাইনাসটা বাইরে বেরিয়ে আসছে"। কিন্তু একটু প্যাঁচ আছে। মাইনাসটা যখন α-র সঙ্গে লাগছে, তখন ওটা হল vector space-এর মাইনাস। আবার বাইরে যখন $-c$ দেখছ, সেই মাইনাসটা হল $\mathbb{R}$-এর মাইনাস। এই দুইধরণের মাইনাসকে সামলানোর জন্য প্রমাণটা একটু ঘুরপথে এগোবে।
এখানে c হল একটা scalar আর α হল একটা vector. বোঝার সুবিধার জন্য আমরা α-কে $\vec{\alpha}$ হিসেবে লিখব।
প্রথমে দেখাব $(-c)\vec{\alpha} = -(c\vec{\alpha})$, মানে $(-c)\vec{\alpha}$ হল $c\vec{\alpha}$-র additive inverse. অর্থাৎ কিনা $(-c)\vec{\alpha} + c\vec{\alpha} = \vec{0}$ দেখালেই হবে।

> $(-c)\vec{\alpha} + c\vec{\alpha} = ((-c) + c)\vec{\alpha} = 0\vec{\alpha}.$

এটা দেখেই ধাঁ করে $\vec{0}$ লিখে বোসো না যেন। $0\vec{\alpha}$ যে $\vec{0}$ হবে সেটা তো সংজ্ঞার মধ্যে ধরা নেই। সুতরাং ওটা প্রমাণের অপেক্ষা রাখে। প্রমাণটা অবশ্য সহজই--

> Now $0\vec{\alpha} + 0\vec{\alpha} = (0 + 0)\vec{\alpha} = 0\vec{\alpha}.$
> So subtracting $0\vec{\alpha}$ from both sides, $0\vec{\alpha} = \vec{0}.$
> Hence $(-c)\vec{\alpha} = -(c\vec{\alpha})$, as required.

এবার দেখাব $-(c\vec{\alpha}) = c(-\vec{\alpha})$. এখানেও একই কায়দা, $c\vec{\alpha} + c(-\vec{\alpha}) = \vec{0}$ দেখালেই হবে।

> Again $c\vec{\alpha} + c(-\vec{\alpha}) = c(\vec{\alpha} + (-\vec{\alpha})) = c\vec{0}.$

এটাকে $\vec{0}$ দেখানো মোটেই শক্ত নয়--

> Now $c\vec{0} + c\vec{0} = c(\vec{0} + \vec{0}) = c\vec{0}.$
> Subtracting $c\vec{0}$ from both sides, $c\vec{0} = \vec{0}.$
> So $c(-\vec{\alpha}) = -(c\vec{\alpha})$, as required.

■

Vector space-এর সংজ্ঞায় vector addition এবং scalar multiplication-এর বিভিন্ন ধর্ম বলা আছে। যে কটা ধর্ম সংজ্ঞায় বলা আছে, তাদের বাইরেও কিন্তু আরও বহু ধর্ম আছে। কিছু উদাহরণ ছিল উপরের অংকটায়, যেমন $0 \cdot \vec{v} = \vec{0}$ হওয়া বা $-1 \cdot \vec{v} = -\vec{v}$ হওয়া, এরকম। এই ধর্মগুলো সংজ্ঞায় সরাসরি বলা না থাকলেও সংজ্ঞার ধর্মগুলো মিলিয়েজুলিয়েই এদের পাওয়া যায়। সাধারণতঃ সংজ্ঞা লেখার সময়ে ধর্মের তালিকাটা যথাসম্ভব ছোটো রাখা হয়, তাতে লিখতে সুবিধা। কিন্তু অনেক সময়ে কিছু ধর্মকে এতই গুরুত্বপূর্ণ মনে করা হয়, যে সেটাকে সংজ্ঞার বাকি ধর্মগুলো থেকে পাওয়া গেলেও, তাও আলাদা উল্লেখ করা হয়। এরকম একটা উদাহরণ দিলাম নীচে।

Example 8: Show that in a vector space defined over $\mathbb{R}$, the vector addition is always commutative, and it follows from the other defined properties.[5] **(2021.CC6.3h)**
SOLUTION:এই অংকে field-টা $\mathbb{R}$ হবার কোনো দরকার নেই।

To show:

$\forall \vec{u}, \vec{v} \in V \quad \vec{u}+\vec{v}=\vec{v}+\vec{u}.$

$\forall \vec{u}, \vec{v}$ Take any $\vec{u}, \vec{v} \in V$.

Then

$$\begin{aligned}(\vec{u}+\vec{v})-(\vec{v}+\vec{u}) &= (\vec{u}+\vec{v})+(-1)(\vec{v}+\vec{u})\\ &= (\vec{u}+\vec{v})+(-1)\vec{v}+(-1)\vec{u}\\ &= (\vec{u}+\vec{v})-\vec{v}-\vec{u}\\ &= \vec{u}+(\vec{v}-\vec{v})-\vec{u}\\ &= \vec{u}+\vec{0}-\vec{u}\\ &= \vec{u}-\vec{u}=\vec{0}.\end{aligned}$$

Hence $\vec{v}+\vec{u}=\vec{u}+\vec{v}$, as required.

■

Exercise 4: উপরের অংকটা দেখে কেউ উৎসাহিত হয়ে বলেছে যে "প্রত্যেক $\vec{v}$-এরই additive inverse আছে" সেই ধর্মটাও সংজ্ঞার বাকি ধর্মগুলো থেকেই প্রমাণ করা যায়, এইভাবে--

> যেকোনো $\vec{v}$-এর জন্যই $(-1)\cdot\vec{v}$-কেই নেব $\vec{v}$-এর additive inverse. কারণ, $\vec{v}+(-1)\cdot\vec{v}=1\cdot\vec{v}+(-1)\cdot\vec{v}=(1+(-1))\cdot\vec{v}=0\cdot\vec{v}=\vec{0}.$

এই যুক্তিতে ভুল কোথায়? ■

DAY 8
Subspace

আমরা গতকাল দেখেছি যে $\mathbb{R}^2$ হল $\mathbb{R}$-এর উপর একটা vector space. এর যেকোনো vector-কেই (x,y) হিসেবে লেখা যায়। তাই $\mathbb{R}^2$-কে একটা গ্রাফ কাগজ বলে ভাবা যায়। এর মধ্যে x-axis-টার দিকে তাকাও (Fig 11). যেসব point-এর বেলায় $y=0$, তাদের নিয়েই তৈরী x-axis-টা। তার মানে x-axis-টা হল $\mathbb{R}^2$-এর একটা subset, যার নাম দিলাম মনে করো X, মানে

$$X=\{(x,0)\ :\ x\in\mathbb{R}\}.$$

লক্ষ করো X-টা কিন্তু নিজেই একটা vector space, কারণ--

- তুমি যদি X-এর মধ্যে দুটো vector নাও $(x_1,0)$ আর $(x_2,0)$, তবে তাদের যোগফল হবে $(x_1+x_2,0)$. যেহেতু দ্বিতীয় স্থানে 0-ই থাকছে, তাই এই যোগফলটাও X-এর মধ্যেই থাকতে বাধ্য।
- একই যুক্তিতে $(x,0)$-কে কোনো সংখ্যা α দিয়ে গুণ করলেও গুণফলটা (মানে $(\alpha x,0)$) আবার X-এর মধ্যেই থাকবে,

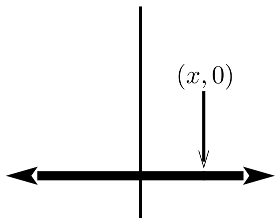


Fig 11

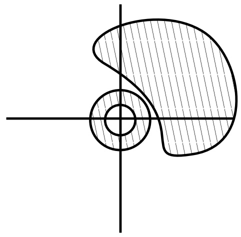


Fig 12

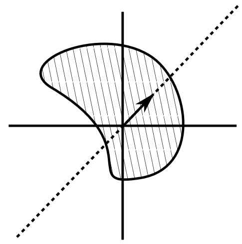


Fig 13

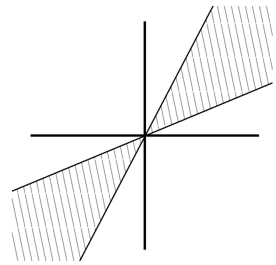


Fig 14

- আর additive identity ছিল $(0,0)$, সেটাও X-এর মধ্যেই আছে।
- বাকি শর্তগুলো (যেমন associativity, commutativity ইত্যাদি) আলাদা করে পরীক্ষা করার দরকার নেই, কারণ ওগুলো তো $\mathbb{R}^2$ থেকেই পেয়ে যাচ্ছি।

এই যে ব্যাপারটা--

> X একাধারে $\mathbb{R}^2$-র একটা subset-ও বটে এবং $\mathbb{R}^2$-র addition আর scalar multiplication ব্যবহার করে নিজেই একটা vector space-ও বটে,

এটাকেই আমরা এক কথায় বলি--

> X হল $\mathbb{R}^2$-র একটা subspace.

সংজ্ঞাটা এইরকম--

> DEFINITION: **Subspace**
>
> Let V be a vector space. If U is a subset of V such that U is also a vector space using the same operations as V, then U is called a **subspace** of V.

যেকোনো vector space-এর মধ্যে সে নিজেই একটা subspace. আবার যেকোনো vector space-এই $\vec{0}$ থাকে, যদি তুমি খালি $\{\vec{0}\}$ নাও, সেটাও কিন্তু একটা subspace হবে (সংজ্ঞা থেকে পরীক্ষা করেই দেখো!)।

8.1 ছবি দিয়ে ভাবা

যদি কোনো vector space-এর একটা subset দেওয়া থাকে, তবে সেটা একটা subspace কিনা, সেটা অনেক সময়ে ছবি দিয়েই বলে দেওয়া যায়। অবশ্যই ছবির প্রসঙ্গ যখন উঠছে, তখন ধরে নিচ্ছি যে, $vector space$-টা $\mathbb{R}^2$ বা $\mathbb{R}^3$. প্রথমে দেখি $\mathbb{R}^2$-এর বেলায় কাজটা কী করে ছবি দেখে করা যায়। প্রথম কথা, $(0,0)$-কে এই subspace-এর মধ্যে থাকতেই হবে। অতএব যদি subset-টা Fig 12-এর মত কিছু একটা হয়, তবে তৎক্ষণাৎ বলে দেওয়া যায় যে, সেটা একটা subspace হতেই পারে না।
$\vec{0}$ থাকা-না-থাকা দেখেই খানিকটা বলা গেল। এবার দেখি closed under scalar multiplication থেকে কী বলা যায়। যদি একটা subspace-এর মধ্যে কোনো $\vec{v}$ থাকে, তবে তার সব multiple-রাও থাকবে। তার মানে ওই $\vec{v}$ বরাবর পুরো লাইনটাও থাকবে। এ থেকেই বলে দেওয়া যায় যে, Fig 13-এর subset-টা subspace নয়, কারণ তীর চিহ্ন দিয়ে দেখানো vector-টা এর ভিতরে আছে, অথচ সেইটা বরাবর পুরো লাইনটা (ড্যাশ্‌ড্যাশ্‌ দিয়ে দেখানো) subset-টার মধ্যে নেই। এবার ধরো Fig 14-এর মত কিছু একটা দিলাম। এটা কি একটা subspace? না, Fig 15 দেখলেই বুঝবে যে, এখানে $\vec{u}$ আর $\vec{v}$ ভিতরে থাকলেও $\vec{u}+\vec{v}$ গিয়ে পড়েছে বাইরে। তাই closed under addition হচ্ছে না। তার মানে যদি origin দিয়ে যাওয়া দুটো লাইন কোনো subspace-এর ভিতরে থাকে, তবে তাদের মাঝের সবকিছুও ওই subspace-এর ভিতরে থাকতে হবে।
এই সবকিছু মিলিয়ে দেখা যায় যে, $\mathbb{R}^2$-এর মধ্যে subspace-রা দেখতে তিন ধরণের (Fig 16)--

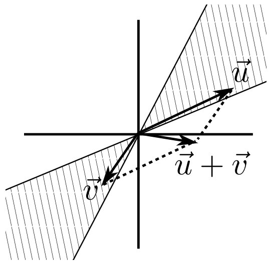


Fig 15

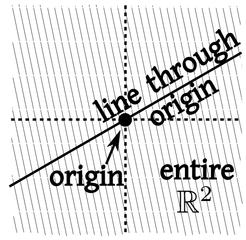


Fig 16

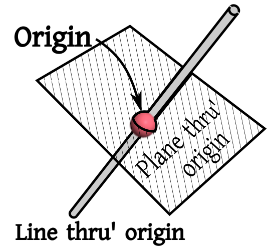


Fig 17

- এক, খালি origin-টা নিয়ে তৈরী একটা singleton set,
- দুই, origin দিয়ে যায়, এরকম যেকোনো সরলরেখা,
- আর তিন, পুরো $\mathbb{R}^2$ নিজে।

একইভাবে $\mathbb{R}^3$-এর ভিতরে subspace-রা চাররকমের (Fig 17)--

- $\{(0,0,0)\}$,
- origin দিয়ে যায় এরকম যেকোনো সরলরেখা,
- origin দিয়ে যায় এরকম যেকোনো plane,
- পুরো $\mathbb{R}^3$ (সেটা আর ছবিতে দেখাতে পারি নি)।

8.2 অংক কষে পরীক্ষা করা

একটা vector space-এর কোনো subset দিয়ে যদি পরীক্ষা করতে বলে সেটা একটা subspace হচ্ছে কিনা, তবে সেই কাজটা খুব সহজে করা যায়, তার জন্য vector space-এর বিশাল সংজ্ঞাটার সবগুলো শর্ত ধরে ধরে পরীক্ষা করতে হয় না, খালি দেখাতে হয় যে subset-টার মধ্যে $\vec{0}$ আছে, এবং subset-টা closed under addition আর closed under scalar multiplication. অর্থাৎ যদি ওর মধ্যে $\vec{u}$ আর $\vec{v}$ থাকে, তবে $\vec{u}+\vec{v}$-ও থাকতে বাধ্য, এবং α কোনো scalar হলে, $\alpha\vec{u}$-ও U-এর মধ্যেই থাকে। এই কথাটা বার বার কাজে লাগবে, তাই একটা theorem হিসেবে লিখে রাখি--

THEOREM

Let V be a vector space over a field F. Let U be any subset of V. Then U is a subspace of V if and only if

1. $\vec{0} \in U$,
2. U is closed under addition, ie,
$$\forall \vec{u}, \vec{v} \in U \quad \vec{u}+\vec{v} \in U,$$
3. U is closed under scalar multiplication, ie,
$$\forall \vec{u} \in U \quad \forall \alpha \in F \quad \alpha\vec{u} \in U.$$

Proof: এটা একটা if and only if জাতীয় জিনিস, তাই প্রমাণটার দুটো অংশ--

> Only if part: If U is a subspace, then U is also a vector space. So the three conditions follow from the definition of vector spaces.

এই অংশটায় কিছুই করতে হল না। অন্য অংশটাও বেশ সংক্ষিপ্ত--

> If part: Due to the given conditions, vector addition and scalar multiplication are well-defined on U.
> Thus, if $\vec{u}, \vec{v} \in U$ and $\alpha \in F$, then $\vec{u} + \vec{v} \in U$ and $\alpha\vec{v} \in U$.
> Also, U has additive identity, since $\vec{0} \in U$.
> All the other properties follow from V.

[Q.E.D]

এই শেষের লাইনটা ভালো করে বোঝা যাক। বাকি শর্তগুলো কী করে V থেকে উত্তরাধিকারসূত্রেই পাওয়া যাচ্ছে? যেমন ধরো যদি commutativity দেখাতে চাও, মানে

$$\forall \vec{u}, \vec{v} \in U \quad \vec{u} + \vec{v} = \vec{v} + \vec{u},$$

তবে এটা লক্ষ করলেই চলবে যে $\vec{u}, \vec{v} \in V$-ও বটে (কারণ $U \subseteq V$)। সুতরাং V-এর মধ্যে $+$-এর commutativity থেকেই পেয়ে যাচ্ছি যে $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ হতে বাধ্য! একইরকম কথা খাটবে বাকি শর্তগুলোর ক্ষেত্রেও।

Example 9: Is the set

$$W = \{(x, y, z) \in \mathbb{R}^3 \ : \ x - 2y + 3z = 0\}$$

a subspace of the real vector space $\mathbb{R}^3$?[3] **(2005.12cp1)**
SOLUTION:

> By a standard result, we know that W is a subspace if and only if $\vec{0} \in W$ and W is closed under vector addition and scalar multiplication.

তিনটে শর্ত আছে, সমাধানেও সেই অনুযায়ী তিনটে অংশ হবে। এককেটা অংশে এককেটা শর্ত পরীক্ষা করে দেখব। একটাও শর্ত পালিত না হলে W-টা subspace হবে না। তিনটেই যদি পালিত হয়, তবে subspace হবে।

> $\vec{0} \in W$:
> Clearly, $\vec{0} \in W$, because $0 - 2 \times 0 + 3 \times 0 = 0$.

ব্যস, প্রথম শর্ত মিলে গেল! এবার দ্বিতীয় শর্ত--

> Closed under addition:
> Let $\vec{u} \equiv (u_1, u_2, u_3)$, $\vec{v} \equiv (v_1, v_2, v_3) \in W$.
> Then $u_1 - 2u_2 + 3u_3 = 0$ and $v_1 - 2v_2 + 3v_3 = 0$.

> Hence $\vec{u}+\vec{v}=(u_1+v_1,u_2+v_2,u_3+v_3)\in W$, because
>
> $(u_1+v_1)-2(u_2+v_2)+3(u_3+v_3)=(u_1-2u_2+3u_3)+(v_1-2v_2+3v_3)=0+0=0.$

পড়ে আছে খালি তৃতীয় শর্ত--

> Closed under scalar multiplication:
>
> Again, if $\alpha\in\mathbb{R}$, then $\alpha\vec{u}=(\alpha u_1,\alpha u_2,\alpha u_3)\in W$, because
>
> $$\alpha u_1-2(\alpha u_2)+3(\alpha u_3)=\alpha(u_1-2u_2+3u_3)=0.$$

যেহেতু তিনটে শর্তই পালিত হল, তাই--

> Hence W is a subspace of $\mathbb{R}^3$.

■

Exercise 5: Let $S=\{(x,y,z)\in\mathbb{R}^3 \ : \ x+2y=z, \quad 2x+3y=z\}$. Prove that S is a subspace of the real vector space $\mathbb{R}^3$.[3] **(2012.3b)**
HINT: আগের অংকের মতই। ■

Exercise 6: Let $W=\{(x,y,z)\in\mathbb{R}^3 \ : \ x-4y+3z=0\}$. Show that W is a subspace of $\mathbb{R}^3$. **(2008.9b)**
HINT: আগের অংকের মতই। ■

এই দুটো অংকেই আমরা $\mathbb{R}^3$-র উপরে কিছু শর্ত চাপিয়ে subspace বানালাম। শর্তগুলোর চেহারা সর্বদাই $ax+by+cz=0$ আকারের ছিল। এই ব্যাপারটা যেকোনো $\mathbb{R}^n$-এর জন্যই হয়। $\mathbb{R}^n$-এর সদস্যরা $(x_1,...,x_n)$-এর মত দেখতে। এদের উপরে তুমি যত খুশি $a_1x_1+\cdots+a_nx_n=0$ জাতীয় শর্ত চাপাও, সর্বদাই কোনো না কোনো subspace পাবে। কিন্তু যদি শর্তটার ডানদিকে শূন্য ছাড়া অন্য কিছু থাকে, তবে কোনোভাবেই আর subspace হবে না (কারণ $\vec{0}$ আর সেই শর্তটা পালন করবে না)। এই ব্যাপারটা মাথায় রেখে নীচের অংকটা করো।

Exercise 7: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). Which of the following is a subspace of $\mathbb{R}^3$?

(A) $2a+3b+5c=0$

(B) $2a+3b+5c=1$

(C) $2a+3b+5c=-1$

(D) None of the above,

where $a,b,c\in\mathbb{R}$.
(2021.CC6.1i) ■

Example 10: Let

$$V_2 = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \;:\; a, b, c \in \mathbb{Q} \text{ (set of rational numbers)} \right\}.$$

Show that V_2 is a subspace of the vector space of 2×2 real matrices with usual matrix addition and scalar multiplication.[3] **(2004.13b)**

SOLUTION: এখানে $\mathbb{Q}$-এর উপরে কাজ করতে হবে, মানে field-টা নিতে হবে $\mathbb{Q}$.

এখানে V_2-র সদস্যরা কীরকম দেখতে বুঝে নাও। এরা প্রত্যেকেই একেকটা 2×2 সাইজের matrix, যাদের entry-গুলো হল rational সংখ্যা। সেই সঙ্গে $(1,2)$-th এবং $(2,1)$-th entry দুটো সমান হতে হবে। যেমন $\begin{bmatrix} \frac{1}{2} & 1 \\ 1 & \\ 0 & \end{bmatrix} \in V_2$ কিন্তু $\begin{bmatrix} \frac{1}{2} & 2 \\ 1 & \\ 0 & \end{bmatrix}, \begin{bmatrix} \pi & 1 \\ 1 & \\ 0 & \end{bmatrix} \notin V_2$.

> Here V_2 is a subset of the vector space of 2×2 real matrices over $\mathbb{Q}$ with usual matrix addition and scalar multiplication.
>
> We know that V_2 is a subspace if and only if the zero matrix is in V_0 and it is closed under addition and scalar multiplication.

এখানেও সমাধানের তিনটে অংশ--

> $\vec{0} \in V_2$: Putting $a = b = c = 0 \in \mathbb{Q}$,
>
> $$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \in V_2.$$
>
> Closed under addition: Let $\begin{bmatrix} a & b \\ b & c \end{bmatrix}, \begin{bmatrix} p & q \\ q & r \end{bmatrix} \in V_2$.
>
> Then $a, b, c, p, q, r \in \mathbb{Q}$. So $a + p, b + q, c + r \in \mathbb{Q}$.
>
> Thus $\begin{bmatrix} a & b \\ b & c \end{bmatrix} + \begin{bmatrix} p & q \\ q & r \end{bmatrix} + \begin{bmatrix} a+p & b+q \\ b+q & c+r \end{bmatrix} \in V_2$.
>
> Closed under scalar multiplication:
>
> Let $\begin{bmatrix} a & b \\ b & c \end{bmatrix} \in V_2$ and $\alpha \in \mathbb{Q}$.
>
> Then $\alpha a, \alpha b, \alpha c \in \mathbb{Q}$.
>
> So $\alpha \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in V_2$.
>
> So V_2 is a subspace of 2×2 real matrices with the usual operations.

■

Exercise 8: If V is the real vector space of all real matrices of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and W is the subset of all matrices of V for which $a+b=0$, then prove that W is a subspace of V.[2] **(2006.13ci)**
HINT: আগের অংকটার মতই। কিন্তু এখানে field-টা $\mathbb{R}$ নিলেই হবে। ■

Exercise 9:

$$V_2 = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \ : \ a+d=0, \ a,b,c,d \in \mathbb{R} \right\}.$$

Show that V_2 is a subspace of the vector space of 2×2 real matrices with usual matrix addition and scalar multiplication. **(2007.12b)**
HINT:
আগের মতই। ■

Vector space-এর সংজ্ঞায় একরাশ শর্ত ছিল। কিন্তু subspace প্রমাণ করার জন্য খালি তিনটে শর্ত দেখালেই চলছিল। এই শর্তগুলো নিয়ে একটু খেলা করা যায়। সেটা করতে গিয়েই নীচের অংকগুলোর জন্ম।

Exercise 10: Let S be a nonempty subset of a vector space V. Show that S is a subspace if and only if it is closed under addition and scalar multiplication.
HINT: এখানেও তিনটে শর্তই দিয়েছে। শেষের দুটো আগের মতই। খালি প্রথম শর্তটা ছিল $\vec{0} \in S$, সেটাকে বদলে করেছে $S \neq \phi$. যদি $\vec{0} \in S$ হয় তবে তো $S \neq \phi$ হবেই। তোমার কাজ হবে উল্টোদিকটা দেখানো, যদি $S \neq \phi$ হয়, তবে $\vec{0} \in S$ হবেই। ■

এই অংকের শর্তগুলোকে আরো সংক্ষেপে লিখলে নীচের অংকের শর্তটা পাওয়া যায়। ভবিষ্যতে আমরা কোনো subset-কে subspace দেখানোর জন্য নীচের অংকের সংক্ষিপ্ত শর্তটাই ব্যবহার করব।

Exercise 11: Let S be a nonempty subset of a vector space V. Show that S is a subspace if and only if

$$\forall \vec{u}, \vec{v} \in S \quad \forall a, b \in F \quad a\vec{u} + b\vec{v} \in S.$$

■

Exercise 12: Let S be a subset of a vector space V such that $\vec{0} \in S$ and $\forall \vec{u}, \vec{v} \in S \quad \vec{u}+\vec{v} \in S$. Show that S need not be a subspace.
HINT: তোমার কাজ হল কোনো একটা counterexample বার করা। মানে এমন কোনো S ভাবা, যার মধ্যে $\vec{0}$ আছে এবং যেটা closed under addition হয়, অথচ subspace না হয়। বুঝতেই পারছ যে, ঝামেলাটা বাঁধছে closed under scalar multiplication না বলায়। সুতরাং এমন S ভাবার চেষ্টা করো, যেটা closed under scalar multiplication না হয়। ■

Exercise 13: Let S be a subset of a real vector space V such that $\vec{0} \in S$ and $\forall \vec{u} \in S \quad \forall \alpha \in \mathbb{R} \quad \alpha\vec{u} \in S$. Show that S need not be a subspace. ■

DAY 9
Linear combination

এবার আমরা কয়েকটা গুরুত্বপূর্ণ জিনিসের সঙ্গে পরিচিত হব--

- linear combination,
- generating set এবং span,
- linear independence,
- basis.

Linear algebra-য় তোমার দক্ষতা প্রচুর পরিমাণে নির্ভর করবে এই চারটে জিনিসে তোমার দখল কতটা, তার উপর। ধরো একটা vector space আছে। সেখান থেকে কয়েকটা vector-এর একটা তালিকা নিলাম। যেমন vector space-টা যদি হয় $\mathbb{R}^3$, তবে আমরা নিলাম ধরো $\vec{v}_1 = (1,2,3), \vec{v}_2 = (0,-1,5),\ \vec{v}_3 = (1,2,3)$ আর $\vec{v}_4 = (2,2,2)$. লক্ষ কর একই vector একাধিক বার নেওয়া যায়, যেমন $\vec{v}_1$ আর $\vec{v}_3$ এখানে একই। এবার যতগুলো $\vec{v}_i$ নিয়েছি ঠিক ততগুলো scalar নেব। আমাদের উদাহরণে field-টা হল $\mathbb{R}$, তাই scalar মানে কিছু real number. চারটে scalar নিলাম ধরো $a_1 = 34, a_2 = -\frac{1}{2}, a_3 = 2$ আর $a_4 = 0$. এবার আমরা scalar-গুলো দিয়ে vector-গুলোকে গুণ করে সবটা যোগ করে দেব। মানে যদি vector-গুলোকে $\vec{v}_1, ..., \vec{v}_n$ আর scalar-গুলোকে $a_1, ..., a_n$ বলি, তবে

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$$

বার করব। আমাদের উদাহরণে

$$\begin{aligned} & 34 \times (1,2,3) - \tfrac{1}{2} \times (0,-1,5) + 2 \times (1,2,3) + 0 \times (2,2,2) \\ = \quad & (36, 72\tfrac{1}{2}, 105\tfrac{1}{2}). \end{aligned}$$

এই কাজটাকে বলে vector-গুলোর একটা **linear combination** নেওয়া। অর্থাৎ যদি কয়েকটা vector-এর একটা তালিকা দিয়ে বলি এদের একটা linear combination নাও, তবে অমনি তুমি তালিকার প্রত্যেকটা vector-এর জন্য যা খুশী একটা করে scalar নেবে, এবং প্রত্যেকটা vector-কে তার জন্য নির্বাচিত scalar দিয়ে গুণ করে যোগ করে দেবে। যদি তোমার তালিকায় খালি একটাই vector থাকে, তবে তার linear combination নেওয়া মানে তাকে কোনো একটা scalar দিয়ে গুণ করা (এক্ষেত্রে আর যোগ করার কিছু নেই)।

এই ব্যাপারটাকে একটু ছবি দিয়ে বুঝে রাখলে পরে অনেক সুবিধা হবে। ছবি আঁকার জন্য আমরা $\mathbb{R}^2$ নিয়ে কাজ করব, যাকে একটা গ্রাফ কাগজ বলে মনে করা যায়। প্রথমে খালি একটা vector নিয়ে শুরু করি, যেমন দেখিয়েছি Fig 18-এ। যদি এটাকে 2 দিয়ে গুণ করি, তবে তীর চিহ্নটা দ্বিগুণ লম্বা হয়ে যাবে। যদি -1 দিয়ে গুণ করি, তবে একই লাইন বরাবর উল্টো হয়ে যাবে। যদি $\frac{1}{2}$ দিয়ে গুণ করি, তবে বেঁটে হয়ে অর্ধেক হয়ে যাবে (Fig 19). আশা করি বুঝতেই পারছ যে, এইভাবে বিভিন্ন scalar দিয়ে গুণ করে ওই লাইন বরাবর যেকোনো vector-ই পাওয়া সম্ভব, এবং ওই লাইনের বাইরের

Fig 18

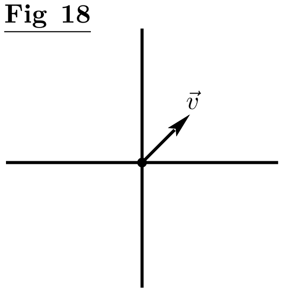
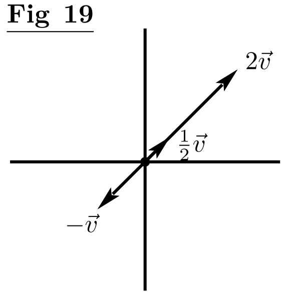


Fig 19



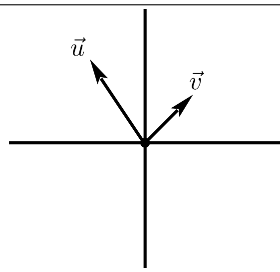


Fig 20

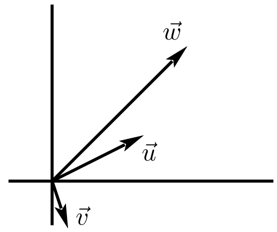


Fig 21

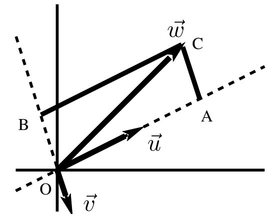


Fig 22

কোনো vector-ই পাওয়া সম্ভব নয়। সুতরাং যদি Fig 20-এর মত দুটো vector $\vec{u}$ আর $\vec{v}$ দিয়ে জিজ্ঞাসা করি যে $\vec{u}$-কে $\vec{v}$-এর linear combination হিসেবে লিখতে পারো কিনা, তবে উত্তর এক কথায় "না", কারণ ওরা একই লাইনে নেই। এই বক্তব্যটাকে অংকের ভাষায় আমরা লিখব

$$\vec{u} \notin \text{span}\{\vec{v}\}.$$

এখানে span$\{\vec{v}\}$ মানে হল $\vec{v}$-এর যাবতীয় linear combination-এর set. ছবিতে span$\{\vec{v}\}$-কে ভাবতে পারো $\vec{v}$ বরাবর লাইনটা হিসেবে।

এবার ধরো দুটো vector দিলাম, $\vec{v}_1$, $\vec{v}_2$. দেখা যাক এদের span-টা কী হয়, মানে এদের দিয়ে কী কী linear combination বানানো যায়। খুব সহজ একটা linear combination হল $\vec{v}_1 + \vec{v}_2$, মানে দুটো scalar-ই এখানে 1 নিয়েছি। তাহলে ছবিতে সেটা একটা parallelogram দিয়ে বোঝানো যায়। যদি নিতাম $2\vec{v}_1 + \vec{v}_2$, তবে প্রথমে $\vec{v}_1$-কে টেনে দ্বিগুণ লম্বা করে তারপরে parallelogram-টা আঁকতে হবে। এই parallelogram আঁকার ব্যাপারটা মাথায় রেখে নীচের অংকটা কর।

Example 11: Fig 21-এ $\vec{u}, \vec{v}$ আর $\vec{w}$, এই তিনটে vector দেখানো আছে। বল তো $\vec{w}$-কে কি $\vec{u}, \vec{v}$-এর linear combination হিসেবে প্রকাশ করা যায়?

SOLUTION: প্রথমে $\vec{u}$ আর $\vec{v}$ বরাবর দুটো লাইন এঁকে নাও। Fig 22-এ ড্যাশ্ ড্যাশ্ দিয়ে দেখিয়েছি। তার পর $\vec{w}$-এর দিয়ে ওদের সঙ্গে parallel করে দুটো লাইন টানো, যেমন মোটা করে দেখিয়েছি। তাহলেই একটা parallelogram তৈরী হবে, $OACB$. এখানে OA রয়েছে $\vec{u}$-এর লাইন বরাবর, মানে $\vec{OA}$ হবে $\vec{u}$-এর একটা multiple. ধরো $\vec{OA} = a\vec{u}$. একইভাবে $\vec{OB}$ হবে $\vec{v}$-র multiple, ধরো $\vec{OB} = b\vec{v}$. (এখানে $b < 0$ হবে কারণ $\vec{OB}$ আর $\vec{v}$ পরস্পরের বিপরীত দিকে মুখ করে রয়েছে)। সুতরাং parallelogram-টা থেকে পাচ্ছি--

$$\vec{w} = \vec{OA} + \vec{OB} = a\vec{u} + b\vec{v}.$$

তার মানে $\vec{w}$-কে $\vec{u}$ আর $\vec{v}$-এর linear combination হিসেবে প্রকাশ করা যাচ্ছে। ■

একটু চিন্তা করলেই বুঝবে যে $\mathbb{R}^2$-র মধ্যে যেভাবেই $\vec{w}$ নিই না কেন, একই কাজ করা যাবে। সুতরাং span$\{\vec{u}, \vec{v}\} = \mathbb{R}^2$. ব্যাপারটাকে এইভাবে ভাবতে পারো $\vec{u}$ আর $\vec{v}$ যেন দুজন পাহারাদার। ওরা দুজনে মিলে যেন $\mathbb{R}^2$-র সর্বত্র নজর রাখছে। এবার তাহলে পুরো জিনিসটা গুছিয়ে লিখে ফেলি।

DEFINITION: Span

If S is any list (possibly empty) of vectors from a vector space V (over a field F) then the set of all linear combinations of vectors from S is called the **span** of S, and written as span(S). Thus

$$\text{span}(S) = \{a_1\vec{v}_1 + \cdots + a_n\vec{v}_n \;:\; a_i \in F,\; \vec{v}_i \in S,\; n \in \mathbb{N}\}.$$

If $S = \{\phi\}$, then we define span$(S) = \{\vec{0}\}$.

যদি আমাদের vector space-টা $\mathbb{R}^2$ বা $\mathbb{R}^3$ হলে span জিনিসটাকে ছবি দিয়ে ভাবা যায়। যদি খালি একটাই nonzero vector নাও, তবে তার span হবে সেই vector বরাবর পুরো লাইনটা। যদি দুটো vector নাও, যারা একই লাইনে নয়, তবে ওদের span হল ওদের মধ্যে দিয়ে যাওয়া plane-টা। $\mathbb{R}^2$-এর মধ্যে এর উপরে আর ওঠা সম্ভব নয়। যদি $\mathbb{R}^3$-র মধ্যে এমন তিনটে vector নাও, যারা সকলেই একই plane-এর নেই, তবে ওদের span হবে পুরো $\mathbb{R}^3$-টাই।

Exercise 14: ধরো F একটা field, আর তার উপরে V কোনো একটা vector space. যদি $S \subseteq V$ হয়, তবে প্রমাণ করো যে, span(S) একটা subspace হতে বাধ্য।
HINT:
সংজ্ঞা থেকেই দেখা যাচ্ছে যে, span$(S) \neq \phi$. আমরা আগেই একটা শর্ত শিখেছিলাম, subspace হবার জন্য, সেটা লাগাও-- যদি $\vec{u}, \vec{v} \in$ span(S) আর $a, b \in F$ হয়, তবে এটা দেখালেই হবে যে $a\vec{u} + b\vec{v} \in$ span(S) হতে বাধ্য। ■

Example 12: Does the vector $(2, 1, 3)$ belong to the subspace of $\mathbb{R}^3$ spanned by the vectors $(1, 2, 3)$ and $(3, 1, 0)$? Justify your answer.[2] **(2019.CC6.3di)**
SOLUTION: এখানে আমাদের দেখতে হবে $(2, 1, 3)$-কে $(1, 2, 3)$ এবং $(3, 1, 0)$-র linear combination হিসেবে লেখা যায় কিনা।

> Let, if possible, $(2, 1, 3) = a(1, 2, 3) + b(3, 1, 0)$ for some $a, b \in \mathbb{R}$.

এখানে "if possible" লিখে নিয়েছি, কারণ এখনও জানি না এরকম a, b পাওয়া সম্ভব কিনা।

> Then $a + 3b = 2$, $2a + b = 1$ and $3a = 3$.
> So $a = 1$, and hence from the second equation $b = -\frac{1}{2}$.
> But this does not satisfy the first equation.
> So $(2, 1, 3) \notin \text{span}\{(1, 2, 3), (3, 1, 0)\}$.

■

একইরকম আরেকটা অংক এবার তোমার জন্য।

Exercise 15: Does the vector $(5, 2, 1)$ belong to the subspace of $\mathbb{R}^3$ spanned by the vectors $(1, 2, 3)$ and $(3, 1, 0)$? Justify your answer. ■

Exercise 16: Show that span(S) is the smallest subspace containing S.
HINT: যেকোনো একটা subspace নাও, ধরো U, যার পেটের মধ্যে S আছে। তবে দেখাও span$(S) \subseteq U$ হবেই। ■

Exercise 17: Prove that the sets $\{(1, 1, 0, 0), (1, 0, 1, 1)\}$ and $\{(2-1, 3, 3), (0, 1, -1, -1)\}$ generate the same vector subspaces of $\mathbb{R}^4$.[5] **(2021.CC6.3e)**
HINT: দেখাও যে, প্রথম set-এর প্রত্যেকেই দ্বিতীয় set-এর span-এর মধ্যে আছে, আবার দ্বিতীয় set-এর প্রত্যেকেও প্রথম set-এর span-এর মধ্যে আছে।
■

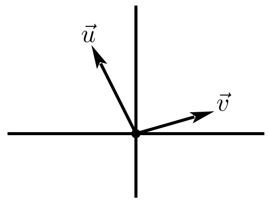


Fig 23

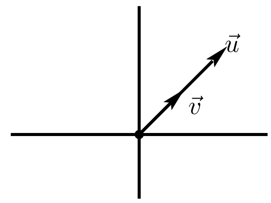


Fig 24

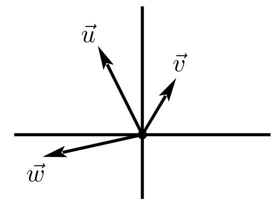


Fig 25

9.1 Spanning or generating set

আমরা জানি যে, square-এর উল্টো হল square root, অর্থাৎ "x-এর square হল y" এই বাক্যটাকেই উল্টে লেখা যায় "y-এর একটা square root হল x"। একইভাবে span-এর উল্টো হল spanning set বা generating set. যদি S-এর span হয় U, তবে U-এর একটা spanning/generating set হল S.

> DEFINITION: **Spanning or generating set**
>
> Let V be a vector space over some field F. Let S be a list of vectors in V, and let U be a subspace of V. If $\text{span}(S) = U$, then we say that S is a **spanning set** or **generating set** of U.

Example 13: এবার Fig 23 দ্যাখো। এখানে $\{\vec{u}, \vec{v}\}$ কি $\mathbb{R}^2$-এর একটা spanning set হবে? যদি Fig 24-এর মত নিতাম তবে?

SOLUTION: Fig 23-এর বেলায় উত্তর হল--হ্যাঁ। কিন্তু Fig 24-এর বেলায়--না। ■

9.1.1 বাড়ালে কী হবে?

Example 14: Fig 25 দ্যাখো। এখানে $\{\vec{u}, \vec{v}\}$ দেখাই যাচ্ছে $\mathbb{R}^2$-এর একটা generating set. মানে আমরা ভাবতে পারি যে $\vec{u}$ আর $\vec{v}$ যেন দুজন পাহারাদার, এবং পুরো $\mathbb{R}^2$-কে পাহারা দেওয়ার জন্য এরা দুজনেই যথেষ্ট। এবার ধরো এদের সঙ্গে আরেকজন পাহারাদার এসে যোগ দিল, $\vec{w}$. প্রশ্ন হল $\{\vec{u}, \vec{v}, \vec{w}\}$-ও কি $\mathbb{R}^2$-র একটা generating set হবে?

SOLUTION: বোঝাই যাচ্ছে যে উত্তর হবে--হ্যাঁ। খালি $\vec{u}, \vec{v}$ দিয়েই তো কাজ চলে যাবে, $\vec{w}$-কে বসিয়ে রাখলেই হবে! যেমন ধরো যদি কোনো $\vec{x} \in \mathbb{R}^2$ দিই, তবে আমরা অমনি তাকে $\vec{x} = a\vec{u} + b\vec{v}$ হিসেবে লিখে ফেলতে পারব। এটাকেই তুমি এইভাবে ভাবো না--

$$\vec{x} = a\vec{u} + b\vec{v} + 0\vec{w}.$$

শূন্য দিয়ে গুণ করা মানেই হল "বসিয়ে রাখা"। ■

সুতরাং নীচের theorem-টা যে সত্যি এ নিয়ে কোনো সংশয়ের অবকাশ নেই।

> THEOREM
>
> Let S be a list of vectors in a vector space V. Suppose that S generates V. Let T be another list of vectors of V such that $S \subseteq T$. Then T must also generate V.

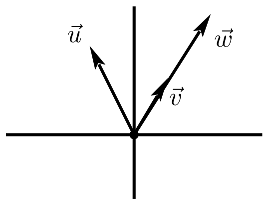


Fig 26

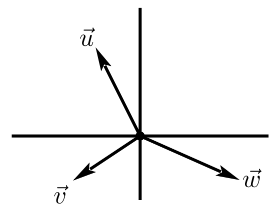


Fig 27

অর্থাৎ তোমার হাতে যদি কিছু vector-এর একটা তালিকা থাকে যারা সবাই মিলে পুরো vector space-টাকে generate করে, তবে সেই তালিকায় আরও বাড়তি vector যোগ করলে নতুন লম্বা তালিকাটাও পুরো vector space-টাকে generate করবে। এবার এর উল্টো দিকটা দেখি। ধরো একটা তালিকা দেওয়া ছিল যেটা দিয়ে পুরো vector space-টা generate করা যায়। তা থেকে vector বাদ দিতে থাকলেও কি অবশিষ্ট তালিকাটা দিয়েও generate করা যাবে? উত্তরটা ভেবে বার করার চেষ্টা কর। আমাদের উত্তরটা নীচে দিলাম, মিলিয়ে নিও।

9.1.2 কমালে কী হবে?

এবার আমাদের উত্তর হল কখনও "হ্যাঁ" আবার কখনও "না", সেটা নীচের উদাহরণটা দেখলেই বুঝবে।

Example 15: Fig 26-এর দিকে তাকাও। আশা করি সন্দেহ নেই যে $\{\vec{u},\vec{v},\vec{w}\}$ দিয়ে পুরো $\mathbb{R}^2$-কে generate করা যায়। যদি আমরা এখান থেকে $\vec{u}$-কে বাদ দিয়ে দিই, তবে পড়ে থাকবে $\{\vec{v},\vec{w}\}$. এটাও দিয়েও কি পুরো $\mathbb{R}^2$-কে generate করা যাবে? আর যদি $\vec{u}$-র বদলে $\vec{v}$-কে বাদ দিতাম?

SOLUTION: যদি $\vec{u}$-কে বাদ দিই, তবে যে দুজন পাহারাদার পড়ে থাকে তাদের দুজনেরই ক্ষমতা একই লাইন বরাবর। ওই লাইনের বাইরে যাবার ক্ষমতা ওদের কারোরই নেই। সুতরাং $\{\vec{v},\vec{w}\}$ দিয়ে পুরো $\mathbb{R}^2$-কে মোটেই generate করা যাবে না। কিন্তু $\vec{v}$-কে বাদ দিলে কোনোই অসুবিধা নেই, কারণ $\vec{v}$ -এর কাজ পুরোটাই $\vec{w}$-ই করে দিতে পারে। একই যুক্তিতে $\vec{v}$-কে রেখে $\vec{w}$-কেও বাদ দেওয়া চলত। ■

লক্ষ কর এখানে $\vec{u}$ ছিল অপরিহার্য, কারণ ওর এমন একটা অবদান আছে যেটা $\vec{v}$ আর $\vec{w}$-কে দিয়ে খালি পাওয়া সম্ভব নয়। কিন্তু $\vec{w}$-এর উপস্থিতিতে $\vec{v}$-র অবদানটা নতুন কিছুই নয়। বাকিদের উপস্থিতিতে একটা vector-এর "নতুন কোনো অবদান থাকা" বলতে কী বোঝাচ্ছি সেটা থেকেই আসে একটা অত্যন্ত গুরুত্বপূর্ণ ধারণা, যার নাম linear independence. আমাদের এবারের আলোচনা সেটা নিয়েই।

9.2 Linear independence

একটা উদাহরণ দিয়ে শুরু করি।

Example 16: Fig 27-এ $\mathbb{R}^2$-এর একটা generating set দেখিয়েছি $\{\vec{u},\vec{v},\vec{w}\}$. এখান থেকে যদি $\vec{u}$-কে বাদ দিই তবেও কি generating set থাকবে? যদি $\vec{u}$-কে রেখে $\vec{v}$-কে বাদ দিতাম? আর যদি $\vec{u},\vec{v}$ দুজনকেই একইসঙ্গে বাদ দেওয়া হয়?

SOLUTION: খালি $\vec{u}$-কে বাদ দিলে কোনোই অসুবিধা নেই। সেটা দেখা যায় এইভাবে, $\vec{u}$ নিজেই হল $\vec{v}$ আর $\vec{w}$-র একটা linear combination. ধরো $\vec{u}=a\vec{v}+b\vec{w}$. সুতরাং $\vec{u},\vec{v},\vec{w}$-এর যেকোনো linear combination যদি নিই $p\vec{u}+q\vec{v}+r\vec{w}$, তবে সেটা খালি $\vec{v}$ আর $\vec{w}$-এর linear combination হিসেবেই লেখা চলে--

$$p\vec{u}+q\vec{v}+r\vec{w}=p(a\vec{v}+b\vec{w})+q\vec{v}+r\vec{w}=(pa+q)\vec{v}+(pb+r)\vec{w}.$$

সুতরাং $\vec{v},\vec{w}$-এর উপস্থিতিতে $\vec{u}$-এর এখানে বাড়তি কোনো অবদান নেই।

একই যুক্তি খাটে খালি $\vec{v}$-কে বাদ দেওয়ার বেলাতেও।
কিন্তু যদি $\vec{u}, \vec{v}$ দুজনকেই একসঙ্গে বাদ দিতে চাও, তবে মুস্কিল হবে। কারণ খালি $\vec{w}$-কে দিয়ে ওদেরকে লেখা যায় না। অর্থাৎ $\vec{u}, \vec{v} \notin \text{span}\{\vec{w}\}$. ■

সুতরাং যদি কিছু vector-এর একটা তালিকা দিই--

$$\vec{v}_1, ..., \vec{v}_n,$$

এবং দিয়ে জানতে চাই-- "এদের মধ্যে থেকে কোনো একটাকে (ধরো $\vec{v}_1$-কে) বাদ দিলে কি span-টা কমে যাবে?" তবে তুমি তৎক্ষণাৎ পরীক্ষা করে দেখবে বাকিদের উপস্থিতিতে $\vec{v}_1$-এর কোনো বাড়তি অবদান আছে কিনা, মানে

$$\vec{v}_1 \notin \text{span}\{\vec{v}_2, ..., \vec{v}_n\}$$

হচ্ছে কিনা। যদি হয়, তবে $\vec{v}_1$-কে বাদ দিলে span-টা কমে যাবে। আর যদি বাড়তি কোনো অবদান না থাকে, তবে নিশ্চিন্তে ওটাকে ঝাঁটার বাড়ি মেরে বিদেয় করে দেওয়া যায়। এইভাবে ঝাঁটার বাড়ি মারতে মারতে এক সময়ে দেখবে যে আর কাউকে বিদায় করা যাচ্ছে না। অর্থাৎ তোমার হাতে এখন এমন কিছু vector-এর তালিকা পড়ে রয়েছে, যেখানে প্রত্যেকেরই নিজস্ব কিছু অবদান আছে, কেউ বাকিদের পয়সায় বসে বসে অন্নধ্বংস করছে না। এরকম vector-দের তালিকাকে বলে linearly independent. অংকের ভাষায়--

DEFINITION: Linearly independent

Let V be a vector space. Let S be a list of vectors in V. Then S is called linearly independent if $\forall \vec{v} \in S \quad \vec{v} \notin \text{span}(S \setminus \{\vec{v}\})$.

এখানে $S \setminus \{\vec{v}\}$ মানে হল S-এর থেকে $\vec{v}$-কে বাদ দিলে যা পড়ে থাকে।
একটা উদাহরণ নেওয়া যাক।

Example 17: এই matrix-টার row-গুলো কি linearly independent

$$\begin{bmatrix} 1 & 2 & 5 \\ 2 & 3 & 0 \\ 1 & 2 & 5 \end{bmatrix}?$$

SOLUTION: এখানে তিনটে vector-এর তালিকা নিয়ে কাজ করছি--

$$\vec{v}_1 = (1, 2, 5), \;\; \vec{v}_2 = (2, 3, 0), \;\; \vec{v}_3 = (1, 2, 5).$$

লক্ষ কর যে এই উদাহরণে $\vec{v}_1$ আর $\vec{v}_3$ আসলে একই। ফলে $\vec{v}_1$-এর উপস্থিতিতে $\vec{v}_3$-কে রাখার কোনো দরকার নেই (বা $\vec{v}_3$-এর উপস্থিতিতে $\vec{v}_1$-কে)। সুতরাং আমরা বলব যে row তিনটে linearly independent নয়, অর্থাৎ ওরা linearly dependent. সংজ্ঞাটা যেভাবে লিখেছিলাম, সেই ভাষায় লিখতে হলে কারণটা লিখব $\vec{v}_3 \in \text{span}\{\vec{v}_1, \vec{v}_2\}$. ■

এই অংকটা থেকে একটা গুরুত্বপূর্ণ শিক্ষা হল--যদি কোনো তালিকায় একই vector একাধিকবার থাকে, তবে সেই তালিকা কোনো দিনই linearly independent হতে পারে না। অবশ্য তালিকায় প্রতিটা vector একবার করে থাকলেই যে তালিকাটা linearly independent হবে, এমন কোনো কথা নেই। যেমন ধরো নীচের অংকটা--

Example 18: আবার আগের অংকের matrix-টা দ্যাখো। বল তো ওর column-গুলো linearly independent কিনা।

SOLUTION: এখানে কাজ হচ্ছে এই তালিকাটা নিয়ে--

$$\vec{v}_1 = (1,2,1), \ \ \vec{v}_2 = (2,3,2), \ \ \vec{v}_3 = (5,0,5).$$

এখানে $\vec{v}_1, \vec{v}_2, \vec{v}_3$ সবাই আলাদা আলাদা। কিন্তু যদি খুব খুঁটিয়ে লক্ষ্য কর তবে দেখবে যে

$$\vec{v}_3 = -15\vec{v}_1 + 10\vec{v}_2.$$

সুতরাং $\vec{v}_3 \in \operatorname{span}\{\vec{v}_1, \vec{v}_2\}$. অতএব column-গুলো linearly independent হচ্ছে না। ■

এখানে এই যে আমরা ফস্ করে লিখে দিলাম "খুঁটিয়ে লক্ষ্য করলে দেখবে", সেটা কিন্তু বড় সহজ কথা নয়। কী করে বুঝলাম যে $\vec{v}_1$-কে -15 আর $\vec{v}_2$-কে 10 দিয়ে গুণ করে যোগ করলেই $\vec{v}_3$ আসবে? যদি একটা vector-এর তালিকা দিয়ে পরীক্ষা করতে বলা হয় সেটা linearly independent কিনা, তবে কি তুমি প্রত্যেকটা vector-কে নিয়ে আলাদা আলাদা করে দেখবে সেটা বাকিদের span-এর মধ্যে আছে কিনা? আমাদের সংজ্ঞা ঠিক সেটাই করতে বলছে, কিন্তু কাজটা মোটেই সহজ নয়। এই সংজ্ঞাটা সেই কারণে ব্যবহার করার পক্ষে একটু অসুবিধাজনক। তাই আমরা অংক কষার সময়ে একটা theorem ব্যবহার করি--

THEOREM

Let V be a vector space. Let S be a list of vectors $\vec{v}_1, ..., \vec{v}_n$ in V. Then S is linearly independent if and only if

$$\forall a_1, ..., a_n \in F \quad (a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0} \implies a_1 = \cdots = a_n = 0).$$

অর্থাৎ তুমি তালিকাটার একটা linear combination নিয়ে সেটা সমান $\vec{0}$ লিখে একটা equation বানাবে--

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}.$$

যদি এই equation-টাকে নাড়াচাড়া করতে করতে এটা দেখাতে পারো যে, প্রতিটা a_i-ই শূন্য হতে বাধ্য, তবেই দেখানো হয়ে গেল যে $\vec{v}_1, ..., \vec{v}_n$ ছিল একটা linearly independent তালিকা।
এমনটা কেন হবে, সেটা বোঝা যাক।
Proof: প্রথমে only if অংশটা দেখি-- ধরো একটা linearly independent তালিকা আছে, $\vec{v}_1, ..., \vec{v}_n$. এদের যা খুশি একটা linear combination নিয়ে $\vec{0}$-র সঙ্গে সমান করে দিলে পাচ্ছি

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}.$$

আমাদের দেখাতে হবে যে, সবগুলো a_i-ই শূন্য হতে বাধ্য। ধরো, তা নয়[1]। তবে অন্ততঃ একটা $a_i \neq 0$ থাকবে। তাহলে সেই $\vec{v}_i$-টাকে বাঁদিকে নিয়ে এসে a_i দিয়ে ভাগ করা যাবে--

$$\vec{v}_i = -\sum_{j\neq i} \tfrac{a_j}{a_i}\vec{v}_j.$$

তার মানে $\vec{v}_i$-কে বাকিদের একটা linear combination হিসেবে লিখে ফেলা গেল, অর্থাৎ $\vec{v}_i$-টা বাকিদের span-এর মধ্যে রয়েছে, যেটা অসম্ভব কারণ তালিকাটা linearly independent বলা ছিল।

[1] বুঝতেই পারছ, আমরা proof by contradiction করছি।

এবার প্রমাণের if অংশটা দেখি। এখানে আমরা শুরু করব এমন কিছু $\vec{v}_1, ..., \vec{v}_n$ নিয়ে যাতে $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}$ হলেই $a_1 = \cdots = a_n = 0$ হয়। দেখাব যে, $\vec{v}_1, ..., \vec{v}_n$ সেক্ষেত্রে linearly independent হতে বাধ্য, মানে কোনো $\vec{v}_i$-ই বাকিদের span-এর মধ্যে নেই। এখানেও proof by contradiction-এর পথে এগোব। শুরু করব উল্টোটা ধরে নিয়ে, মানে এমন কোনো $\vec{v}_i$ আছে ধরে নিয়ে, যে পরের পয়সায় বসে খাচ্ছে, অর্থাৎ বাকিদের span-এর মধ্যে আছে--

$$\vec{v}_i = \sum_{j \neq i} a_j \vec{v}_j,$$

যেখানে a_j-রা সবাই scalar. কিন্তু এটাকেই সাজিয়ে লিখলে হবে

$$\sum_{j=1}^{n} a_j \vec{v}_j = \vec{0},$$

যেখানে আমরা $\vec{v}_i$-এর coefficient নিয়েছি $a_i = -1$. গুলিয়ে গেল? আচ্ছা, একটা উদাহরণ নেওয়া যাক--

> ধরো $\vec{v}_1, \vec{v}_2, \vec{v}_3$ ছিল, যেখানে $\vec{v}_2 = a_1\vec{v}_1 + a_3\vec{v}_3$. সেটাকে আমরা সাজিয়ে লিখতে পারি $a_1\vec{v}_1 + (-1)\vec{v}_2 + a_3\vec{v}_3 = \vec{0}$. দেখলে $\vec{v}_2$-র coefficient কীভাবে -1 এসে গেল?

যেহেতু এখানে $a_i = -1 \neq 0$, তাই এখান থেকে একটা contradiction পেয়ে যাচ্ছ। **[Q.E.D]**

এই theorem-টা কাজে লাগবে নীচের অংকটায়।

Example 19: Prove that the set of vectors $\{(1,2,2),(2,1,2),(2,2,1)\}$ is linearly independent in $\mathbb{R}^3$.[3] **(2019.CC6.3dii)**

SOLUTION:

> We shall show that if
>
> $$a(1,2,2) + b(2,1,2) + c(2,2,1) = (0,0,0),$$
>
> for $a, b, c \in \mathbb{R}$, then $a = b = c = 0$.
> We have
>
> $$\begin{aligned} a + 2b + 2c &= 0, \\ 2a + b + 2c &= 0, \\ 2a + 2b + c &= 0. \end{aligned}$$
>
> Adding we have $5(a+b+c) = 0$. So $c = -a - b$.
>
> Hence from the last equation $2(a+b) - (a+b) = 0$, or $a + b = 0$. So $c = 0$.
>
> So, from the first equation, $a + 2b = 0$. Since $a + b = 0$, this means $b = 0$ and hence $a = 0$.
>
> Thus, the three given vectors form an linearly independent set.

■

Exercise 18: ধরো $\{\vec{v}_1, ..., \vec{v}_n\}$ হল linearly independent. যদি

$$\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = \beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n$$

হয়, তবে দেখাও যে, $\forall i \quad \alpha_i = \beta_i$ হবেই। ■

Exercise 19: ধরো $\{\vec{v}_1, ..., \vec{v}_n\}$ এমন একটা set, যাতে যখনই

$$\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = \beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n$$

হয়, তখনই $\forall i \quad \alpha_i = \beta_i$ হবেই। প্রমাণ করো যে, set-টা linearly independent হতে বাধ্য। ■

Theorem-টা দিয়ে অংক কষা সহজ বটে, কিন্তু সমস্যা হল ওটা কেমন যেন যান্ত্রিক দেখতে। ওতে বোঝা যায় না যে ব্যাপারটা কী হচ্ছে, যেটা বোঝা যাচ্ছিল সংজ্ঞাটা থেকে--independence মানে কেউ কারো পয়সায় বসে খায় না, সবারই নিজস্ব একটা বাড়তি অবদান আছে। সেটা অল্প কয়েকটা vector-র ক্ষেত্রে সাধারণতঃ একটু তীর চিহ্ন দিয়ে চিন্তা করেই ভালো বোঝা যায়।

9.3 তীর চিহ্ন দিয়ে ভাবা

যদি তোমার তালিকায় একটাই vector থাকে, তবে সেটা nonnull হলেই linearly independent হবে। এখানে nonnull হওয়াটা দরকার, নইলে তীর চিহ্ন ভাবতেই পারবে না। যদি দুটো vector থাকে, তবে linearly independent হবার জন্য দেখতে হবে যেন ওরা কেউ কারো multiple হয়ে না যায়, মানে যদি vector-গুলো তীর চিহ্ন দিয়ে ভাবো, তবে তীর চিহ্ন দুটো যেন একই লাইন বরাবর না হয়ে যায়। যদি তিনটে vector-এর তালিকা হয়, তবে দেখতে হবে যেন ওরা সকলেই একই plane-এ না থাকে (অবশ্যই plane কল্পনা করার জন্য আমি vector space-টাকে $\mathbb{R}^3$ নিচ্ছি)। এর চেয়ে বেশী লম্বা তালিকা হলে অবশ্য তীর চিহ্ন দিয়ে ভাবা কঠিন।
কিন্তু লম্বা তালিকার বেলাতেও চিন্তা করার কায়দা আছে। এরকম দুটো জিনিস দিলাম নীচের দুটো অংকে। এই দুটো অংকই কাজে লাগবে পরে।
প্রথম অংকটাকে বলা যায় যেন ধাপে ধাপে linear independence দেখানো--

Exercise 20: Let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_3\}$ be a list of vectors such that

$$\vec{v}_1 \neq \vec{0}, \quad \vec{v}_2 \notin \text{span}\{\vec{v}_1\}, \quad \vec{v}_3 \notin \text{span}\{\vec{v}_1, \vec{v}_2\}, ...$$

Then show that $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ is an independent set. ■

যদি দুটো linearly independent তালিকা দিই, তবে ওদের জুড়ে দিলে কি ফের একটা linearly independent তালিকা পাবে? উত্তরটা যে "না", সেটা বোঝাই যায়। যেমন প্রথম তালিকাটা $\{(1,0)\}$ আর দ্বিতীয়টা $\{(2,0)\}$ হলে বুঝতেই পারছ যে $\{(1,0),(2,0)\}$ মোটেই linearly independent হবে না। এখানে সমস্যা হয়েছিল যে দ্বিতীয় vector-টা আসলে প্রথমটারই multiple ছিল। কিন্তু এছাড়াও সমস্যা সম্ভব। তাই নিয়েই নীচের সামান্য ঠকানো প্রশ্নটা--

Exercise 21: Let $A = \{\vec{u}_1, ..., \vec{u}_m\}$ and $B = \{\vec{v}_1, ..., \vec{v}_n\}$ be two independent sets. Also no $\vec{u}_i$ belongs to span(B), and no $\vec{v}_j$ belongs to span(A). Then is it true that $A \cup B$ must be linearly independent? ■

কোনো কোনো ক্ষেত্রে কিছু vector-এর একটা তালিকা যে linearly independent নয়, সেটা চোখে দেখেই বল দেওয়া যায়। এরকম একটা ক্ষেত্র আমরা আগেই দেখেছি--যখন তালিকায় একই vector একাধিক বার আসে। যদি একই vector-এর কোনো multiple-ও আসে, তবেও linearly independent হবার আশায় ছাই। তেমনি আরেকটা বিশেষ ক্ষেত্র নিয়ে নীচের অংকটা, যদি তালিকায় $\vec{0}$ থাকে, তবে আর সেই তালিকার linearly independent হয়ে ওঠা হল না।

Example 20: Prove that a set of vectors containing the null vector in a vector space $V(F)$ is

linearly dependent.[2] **(2006)**
SOLUTION: এখানে $V(F)$ বলতে বোঝানো হচ্ছে যে F নামক field-এর উপরে V নামক vector space. এই notation-টা মোটেও প্রচলিত নয়। এই অংকের বাইরে এটা ব্যবহার না করাই ভালো।

> Let $S \subseteq V(F)$ be such that $\vec{0} \in S$.
> Shall show that S is linearly dependent, ie, there is a subset $\{\vec{v}_1, ..., \vec{v}_n\} \subseteq S$ and scalars $a_1, ..., a_n \in F$, not all zero, such that
> $$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}.$$
> We choose the set $\{\vec{0}\} \subseteq S$ and choose $a_1 = 1 \in F$.
> Then $a_1 \neq 0$, but $a_1\vec{0} = \vec{0}$.
> So S is linearly dependent, as required.

■

9.3.1 বাড়ালে বা কমালে কী হয়?

আমরা আগেই বলেছিলাম যে একটা generating set-এ আরো বাড়তি vector ঢোকালে সেটা generating set-ই থাকে, কিন্তু কিছু vector বাদ দিয়ে দিলে generating set থাকতেও পারে, নাও পারে। এর ঠিক উল্টোটা হয় linearly independent set-দের বেলায়। সেটা সহজে বোঝা যাবে সংজ্ঞা থেকেই। যদি প্রত্যেকটা vector-এরই নিজস্ব অবদান থাকে, তবে বাকিদের সংখ্যা কমে গেলে তো সেই অবদান আরো বেশী করে ফুটে উঠবে! কিন্তু যদি নতুন নতুন vector যোগ করতে থাকো তবে সেই অবদান চাপা পড়ে যেতে পারে, হয়তো একই অবদান-ওয়ালা আরেকজন এসে পড়ল, তখন আর আগের জন অপরিহার্য রইল না। যেমন ধরো নীচের উদাহরণটা--

Example 21: $\vec{v}_1 = (1,2)$ আর $\vec{v}_2 = (0,1)$ কি linearly independent? যদি এই তালিকার সঙ্গে $\vec{v}_3 = (2,4)$-কে জুড়ে দিই তবে?
SOLUTION: $\vec{v}_1$ আর $\vec{v}_2$ কেউ কারো multiple (গুণিতক) নয়, অতএব $\vec{v}_1, \vec{v}_2$ অবশ্যই linearly independent. কিন্তু $\vec{v}_3 = 2\vec{v}_1$. তাই $\vec{v}_1, \vec{v}_2, \vec{v}_3$ একসঙ্গে আর linearly independent রইল না। ■

DAY 10
Basis

10.1 জিনিসটা কী?

আমরা দেখলাম যে, কোনো generating set-এর superset-রাও generating set হয় (যদিও subset-দের বেলায় সে গ্যারান্টি নেই)। আবার কোনো linearly independent set-এর subset-রাও linearly independent হতে বাধ্য (যদিও এখানে আবার superset-দের বেলায় গ্যারান্টিটা নেই)। এই প্রসঙ্গে একটা উপমা মনে পড়ল--কোনো nonnegative সংখ্যার থেকে বড় সংখ্যারাও nonnegative হয়, আর কোনো nonpositive সংখ্যার চেয়ে ছোটো সংখ্যারাও nonpositive-ই হতে বাধ্য। তাই যদি আমরা $\mathbb{R}$-কে একটা লাইন হিসেবে কল্পনা করি, তবে তার ডানদিকটা nonnegative-দের কবলে, আর বাঁদিকটা nonpositive-দের। এই দুই অংশের মোলাকাত হয় যেখানে, তাকে বলা যায় $\mathbb{R}$-এর সবচেয়ে গুরুত্বপূর্ণ সদস্য, 0. এই উপমার সঙ্গে তাল রেখে আমরা যদি generating set আর linearly independent set-দের একটা লাইন টেনে ভাবি

তবে পাব Fig 28-এর মত একটা ছবি। Generating set-রা রয়েছে ডানদিকের অংশ জুড়ে, আর linearly independent set-রা রয়েছে বাঁদিকের অংশে। এই দুই অংশের সন্ধিস্থলে রয়েছে সেই সব set, যারা একই সঙ্গে generating set এবং linearly independent set. বুঝতেই পারছ এরকম বাঘা বাঘা দুটো শর্ত যারা একই সঙ্গে পালন করে, সেই সব set-রা খুবই কাজের জিনিস। এদের বলে **basis**. এর সংজ্ঞাটাই গুছিয়ে লিখতে বলেছে নীচের অংকে। এই সংজ্ঞাটা যেন কখনও ভুলো না।

Example 22: Define basis of a vector space.[1] **(2004.13a)**

SOLUTION:

> DEFINITION: **Basis**
>
> By a <u>basis</u> of a vector space we mean a subset of the vector space which is both linearly independent and a generating set. The empty set is defined to be the basis of the trivial vector space, $\{\vec{0}\}$.

■

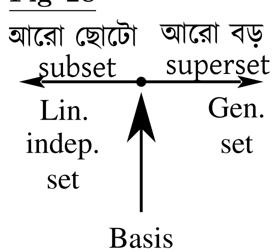
একদম শেষে empty set নিয়ে যে কথাটা বলেছি সেটা একটু পরেই বোঝাচ্ছি। আপাততঃ একবার Fig 28-এর দিকে তাকালেই বুঝবে যে--

- একটা basis-এর চেয়ে বড় linearly independent set সম্ভব নয়। যদি একটা basis-এ আর একটাও vector ঢোকাও, তবে সেই নতুন set-টা আর linearly independent থাকবে না। তাই আমরা বলি যে একটা basis হল একটা maximal linearly independent set. ঠিক যেমন 0 হল সবচেয়ে বড় nonpositive সংখ্যা।

- বিপরীতপক্ষে, একটা basis আবার minimal generating set-ও বটে। একটাও vector বাদ দিলে set-টা আর generating set থাকবে না। ঠিক যেমন 0 হল সবচেয়ে ছোটো nonnegative সংখ্যা।

এই দুটো তথ্যই আমরা ছবি থেকে অনায়াসে দেখতে পাচ্ছি বটে, কিন্তু ভুলে যেও না যে, ছবিটার পিছনে আমরা এখনো কোনো অংকের যুক্তি দেখাই নি। ওটা আপাততঃ নিছক একটা উপমার ভিত্তিতে আঁকা কাল্পনিক ছবি ছাড়া আর কিছুই নয়। $\mathbb{R}$-এর বেলায় আমরা জানি যে nonnegative আর nonpositive-দের মোলাকাত হয় 0-তে এসে। কিন্তু কী করে জানলাম যে linearly independent set আর generating set-দের মধ্যে আদৌ কোথাও মোলাকাত হয়? অর্থাৎ, কী করে দাবী করছি যে, এমন set সম্ভব যেটা একই সঙ্গে linearly independent এবং generating? যেকোনো গুরুতর প্রশ্ন আলোচনা করার আগে কয়েকটা উদাহরণ দেখে নিলে মাথাটা ভালো কাজ করে। তাই কয়েকটা উদাহরণ দেখে নিই।

Fig 28



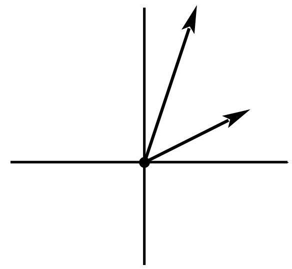


Fig 29

Example 23: Fig 29-এ যে দুটো vector দেখিয়েছি, ওরা মিলে কি $\mathbb{R}^2$-এর একটা basis হয়?

SOLUTION: হ্যাঁ কারণ ওরা মিলে পুরো $\mathbb{R}^2$-কেই পাহারা দিতে পারে, এবং এই কাজের জন্য দুজনেই অপরিহার্য। এক জনকেও বাদ দাও, অমনি আর generating থাকবে না। বাড়তি এক জনকেও ঢোকাও, অমনি আর linearly independent থাকবে না। ■

এক্ষুণি basis-এর যে সংজ্ঞাটা দিলাম তার শেষ লাইনে trivial vector space-দের নিয়ে একটা মন্তব্য ছিল, এবার সেটা বোঝা যাক। অংক কষে প্রমাণ করে দেওয়া যায় যে, যেকোনো nontrivial vector space-এর মধ্যেই basis থাকতে বাধ্য! সে বিষয়ে আমরা একটু পরেই আলোচনা করব। যদি trivial vector space-এরও basis থাকে, তবে সবটা মিলিয়ে বেশ সুন্দরভাবে বলা যায় যে, "সব vector space-এরই basis আছে"। এখন trivial vector space হল $V = \{\vec{0}\}$, মানে এমন একটা vector space যার মধ্যে খালি একটাই vector আছে, $\vec{0}$. এর basis হবার জন্য প্রথমে এর একটা subset নিতে হবে। এখানে subset বলতে খালি দুটো, ϕ আর $\{\vec{0}\}$. এর মধ্যে দ্বিতীয়টা linearly independent নয়, কারণ এর মধ্যে $\vec{0}$ আছে! সুতরাং পড়ে রইল ϕ. কিন্তু ϕ কি linearly independent? হ্যাঁ, কারণ linearly dependent হতে হলে ওর মধ্যে এমন একটা vector থাকা দরকার যেটা বাকিদের span-এ আছে। কিন্তু ϕ-এর মধ্যে তো কোনো vector-ই নেই। তাই ϕ-কে linearly independent বলতে কেউ আপত্তি তুলতে পারে না। এটা অনেকটা যেন বিরুদ্ধপ্রমাণাভাবে আসামীর খালাস পেয়ে যাওয়ার মত, অংকের জগতে একে বলে vacuously true. এবার প্রশ্ন ওঠে ϕ-এর span কী হবে। এক দিক দিয়ে দেখলে span(ϕ) আবার ϕ-ই হওয়া উচিত। তাহলে তো আর V-এর কোনো basis থাকে না। সেই কারণে অনেকে সংজ্ঞার মধ্যে span(ϕ) $= \{\vec{0}\}$ ধরে নেয়, যেমন এই বইতে আমরা ধরেছি। এটা কতকটা যেন $0! = 1$ হওয়ার মত ব্যাপার। অবশ্য অনেকে span(ϕ) $= \{\vec{0}\}$ ধরে না, তারা বলে যে trivial vector space-এর কোনো basis নেই। তবে trivial vector space নিয়ে ভীষণ কিছু মাথা ঘামানোর প্রয়োজন পড়ে না, তাই এই মতপার্থক্যটা কোনো গুরুতর সমস্যা নয়।
এবার আমরা basis-দের একটা অদ্ভুত ধর্মের কথা বলে নিই।

Example 24: $\mathbb{R}^2$-এর একটা basis আমরা দেখেছি Fig 29-এ। এবার এমন একটা basis দিতে পারো, যার সাইজ 2 নয়?

SOLUTION: আমাদের vector space-টা যেহেতু trivial নয়, তাই এর basis মানেই nonempty set, সুতরাং সাইজ 0 হতে পারে না। দেখি 1 সাইজের একটা basis পাই কিনা। ধরো $\{\vec{v}\}$ একটা basis হল, তার মানে $\vec{v}$ এমন একজন পাহারাদার যে পুরো $\mathbb{R}^2$-কে পাহারা দিতে পারে। কিন্তু একা $\vec{v}$-এর কেরামতি তো খালি ওর নিজের লাইনটুকু বরাবর। সুতরাং পুরো $\mathbb{R}^2$-র দায়িত্ব নেওয়া খালি $\vec{v}$-এর কম্মো নয়! আচ্ছা ওকে যদি একটা সহকর্মী দেওয়া যায়, ধরো $\vec{u}$, যে $\vec{v}$-এর লাইন বরাবর নয়? Parallelogram-এর ছবি কল্পনা করেই আমরা জানি যে $\{\vec{u}, \vec{v}\}$ একটা generating set, এবং linearly independent-ও বটে, মানে একটা basis. কিন্তু আমরা তো 2 সাইজের basis চাই না। তাই দেখি আরো কিছু vector এর মধ্যে ঢোকানো যায় কিনা। এমন vector চাই যেটা span$\{\vec{u}, \vec{v}\}$-এর বাইরে থাকবে (নইলে নতুন set-টা linearly independent থাকবে না)। সমস্যা হল span$\{\vec{u}, \vec{v}\} = \mathbb{R}^2$, সুতরাং এমন vector পাওয়াই যাবে না। অতএব $\mathbb{R}^2$-এর যেকোনো basis-এরই সাইজ হবে 2. ■

ঠিক একইরকম ব্যাপার যেকোনো nontrivial vector space-এর ক্ষেত্রেই খাটে। ধরো V যেকোনো একটা nontrivial vector space, আর $B_1, B_2 \subseteq V$ দুজনেই V-এর basis. তাহলে B_1 আর B_2-র সাইজ সমান হতে বাধ্য! সুতরাং

সাইজটা কোনো basis-এর নিজস্ব সম্পত্তি নয়, ওটা vector space-টার বৈশিষ্ট্য। এই সাইজটাকে বলে vector space-টার dimension. সংজ্ঞাটা এইরকম--

> **DEFINITION: Dimension**
>
> Any two bases of the same vector space has the same size(possibly infinite). This size is called the **dimension** of the vector space.

এই সব জিনিস কী করে প্রমাণ করে সে আলোচনায় আমরা পরে যাব। প্রথমে কয়েকটা অংক কষে পুরো ধারণাটা বুঝে নিই।

10.2 হাতে-কলমে

Example 25: Is the set

$$W = \{(x, y, z) \in \mathbb{R}^3 \ : \ x - 2y + 3z = 0\}$$

a subspace of the real vector space $\mathbb{R}^3$? If so, find the dimension of this subspace.[5] **(2005.12cp2)**
SOLUTION: প্রথম অংশের সমাধান তো আগেই করেছি Example 9-এ। এখানে দ্বিতীয় অংশটা করছি।

> Second part: Here $(x, y, z) \in W$ iff $x - 2y + 3z = 0$,

এখানে (x, y, z)-এর মধ্যে তিনটে অক্ষর রয়েছে, কিন্তু $x - 2y + 3z = 0$ হওয়ার সুবাদে, এদের মধ্যে যেকোনো দুটো রাখলেই চলে। ধরো আমরা y আর z রাখলাম। এবার y আর z-এর অংশদুটোকে আলাদা করে লিখব--

> ie, iff (x, y, z) is of the form $(2y - 3z, y, z) = y(2, 1, 0) + z(-3, 0, 1)$ for any $y, z \in \mathbb{R}$.

তার মানে $(2, 1, 0)$ এবং $(-3, 0, 1)$-এর যাবতীয় linear combination দিয়েই W তৈরী। তাই--

> Thus $\{(2, 1, 0),\ (-3, 0, 1)\}$ generates W.
>
> Since these two vectors are not multiples of one another, this set is also linearly independent, and hence a basis of W.
>
> So the required dimension of W is the size of this basis which is 2.

■

Exercise 22: Find the dimension of the subspace S of $\mathbb{R}^3$ where $S = \{(x, y, z) \in \mathbb{R}^3 \ : \ 2x + y - z = 0\}$.[3] **(2022.CC6.3bi)**
HINT:
আগের অংকটার মতই। ■

Exercise 23: Find a basis and dimension of the subspace W of $\mathbb{R}^3$ where

$$W = \{(x, y, z) \ : \ x + y + z = 0\}.$$

[2+1] **(2009.9b)**
HINT:
আগের অংকটার মতই। ■

Exercise 24: Let $S = \{(x, y, z) \in \mathbb{R}^3 \ : \ x + 2y = z, \quad 2x + 3y = z\}$. Prove that S is a subspace of the real vector space $\mathbb{R}^3$. Also find a basis of S and the dimension of S.[5] **(2012.3b)**
HINT:
আগের অংকের মতই। ■

Exercise 25: Let $W = \{(x, y, z) \in \mathbb{R}^3 \ : \ x - 4y + 3z = 0\}$. Show that W is a subspace of $\mathbb{R}^3$. Find the dimension of W. **(2008.9b)**
HINT:
আগের অংকের মতই। ■

Example 26: Let

$$V_2 = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \ : \ a, b, c \in \mathbb{Q} \text{ (set of rational numbers)} \right\}.$$

Show that V_2 is a subspace of the vector space of 2×2 real matrices with usual matrix addition and scalar multiplication. Write a basis of V_2 and find dimension of V_2.[5] **(2004.13b)**
SOLUTION: এখানে field-টা নিতে হবে $\mathbb{Q}$. এই অংকের প্রথম অংশটা করেছিলাম Example 10-এ। এবার দ্বিতীয় অংশটা। V_2-র প্রতিটা সদস্য হল একটা matrix, যাকে বোঝাতে তিনটে অক্ষর a, b আর c ব্যবহার করা হয়েছে। আমরা এই তিনজনের অবদানকে প্রথমে আলাদা করে নেব--

> Second part: A matrix is in V_2 iff it is of the form
>
> $$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$
>
> for some $a, b, c \in \mathbb{Q}$.
> So V_2 is generated by the set
>
> $$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$
>
> This set is also linearly independent.

[[Because:

If $p, q, r \in \mathbb{Q}$ such that

$$p\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + q\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + r\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

then

$$\begin{bmatrix} p & q \\ q & r \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

or $p = q = r = 0$.

]]

Thus S is a basis of V_2, and the dimension of V_2 is $|S| = 3$.

■

Exercise 26: If V is the real vector space of all real matrices of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and W is the subset of all matrices of V for which $a + b = 0$, then prove that W is a subspace of V and find a basis of W.[3] **(2006.13ci)**
HINT:
আগের অংকটার মতই। ■

Exercise 27:

$$V_2 = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \;:\; a + d = 0,\; a, b, c, d \in \mathbb{R} \right\}.$$

Show that V_2 is a subspace of the vector space of 2×2 real matrices with usual matrix addition and scalar multiplication. Write a basis of V_2 and find the dimension of V_2. **(2007.12b)**
HINT:
আগের মতই। ■

10.3 সহজে চিন্তা করার কায়দা

Basis এবং dimension আমাদের বারবার কাজে লাগবে, তাই এদের নিয়ে সহজে চিন্তা করতে পারাটা খুবই জরুরী। জটিল অংক কষে প্রমাণ করতে পারা এক জিনিস, আর চট্ করে গন্ধ শুঁকে বুঝতে পারা আরেক জিনিস। এখানে আমরা এরকম দুটো "গন্ধ শুঁকে" কাজ করার কায়দা বলব।

10.3.1 Basis-রা যেন গ্রাফ কাগজের axis

আমরা বলেছি যে basis হল এমন একটা set যেটা একাধারে linearly independent এবং generating. এই রকম দুটো জবরদস্ত্ শর্ত একসঙ্গে পালিত হচ্ছে মানে basis একটা খুব শক্তিশালী জিনিস, সেটা আন্দাজ করতে পারছ। কিন্তু এই বর্ণনা থেকে বোঝা যাচ্ছে না basis দিয়ে কাজটা কী হয়। আসলে basis-রা আমাদের খুব পরিচিত জিনিস, আমরা স্কুল থেকেই এদের দেখে আসছি, খালি সামান্য অন্য চেহারায়। সেইটা একবার বুঝে গেলেই দেখবে basis-রা শুধু কাজের জিনিসই নয়, কাছের জিনিসও বটে!
একটা উদাহরণ নিয়ে শুরু করি।

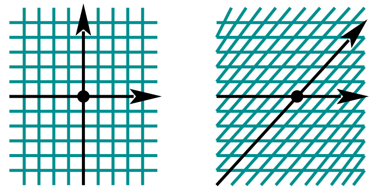


Fig 30

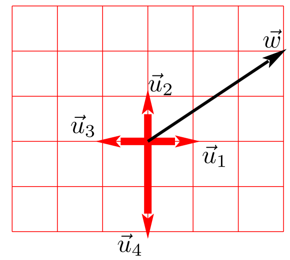


Fig 31

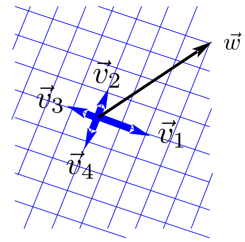


Fig 32

Example 27: ধরো $\mathbb{R}^2$. এটাকে তুমি একটা খাতার পাতা বলে ভাবতে পারো। যদি এর উপর একটা কোনো বিন্দু দিই, তবে তার অবস্থান কী করে বোঝাবে? এর একটা কায়দা হল খাতার পাতাটাকে গ্রাফ কাগজ বলে ভাবা। তা হলে যেকোনো বিন্দুর অবস্থানকে আমরা (x, y) আকারে প্রকাশ করতে পারব। "গ্রাফ কাগজ বলে ভাবা" বলতে কী বোঝাচ্ছি? এর মানে তিনটে জিনিস--

1. দুটো লাইন টানা, x-axis আর y-axis.
2. Axis দুটোর কোন দিকটাকে positive ধরব সেটা ঠিক করা,
3. প্রত্যেকটা axis-এর জন্য একটা করে unit (একক দৈর্ঘ্য) ঠিক করা।

এই তিনটে জিনিসকে মিলে পেলাম একটা coordinate system. এটা নানাভাবে করা যায়। সেই অনুযায়ী গ্রাফ কাগজের ছকগুলো বিভিন্ন হবে। Fig 30 দ্যাখো। লক্ষ্য কর যে axis-দুটো পরস্পরের সঙ্গে perpendicular না হলেও কোনো অসুবিধা নেই। এবার আমরা দেখব যে একটা coordinate system আসলে একটা basis ছাড়া আর কিছুই নয়। সেটা বোঝার জন্য দুটো তীর চিহ্ন আঁকব, একটা হল $(0, 0)$ থেকে $(1, 0)$ অবধি, আর অন্যটা $(0, 0)$ থেকে $(0, 1)$ অবধি। এই তীরদুটো দেওয়া থাকলেই পুরো coordinate system-টা বার করে ফেলতে পারব। ওদের বরাবর লাইন দুটোই হবে axis. তীর দুটোর কোন দিকে মুখ তা থেকে বুঝব কোনদিকটা positive, আর ওদের দৈর্ঘ্যই বলে দিচ্ছে একক দৈর্ঘ্যদুটো কী নিয়েছি। তার মানে basis আর coordinate system আসলে সমার্থক। একটা coordinate system-এ axis দুটো একই লাইন হয়ে গেলে চলবে না, ঠিক সেটাই ছিল basis-এর বেলায় linear independence-এর শর্ত। ■

এখানে অবশ্য একটা ছোট্টো সমস্যা আছে। Axis-গুলোর একটাকে বলি x-axis, আর অন্যটা হল y-axis. আমরা যদি কোনো বিন্দুর অবস্থান লিখি $(3, 4)$ তার মানে 3 হল x-axis বরাবর দূরত্ব, আর 4 হল y-axis বরাবর। কিন্তু basis তো খালি একটা set মাত্র, তার মধ্যে তো কে আগে, কে পরে তার কোনো মূল্য নেই। যদি $\{\vec{u}, \vec{v}\}$ একটা basis হয়, তবে $\{\vec{v}, \vec{u}\}$-ও একই basis. কিন্তু coordinate system হিসেবে এরা তো একই জিনিস নয়! সেই কারণে আমরা সাধারণতঃ খালি basis নিয়ে কাজ না করে **ordered basis** নিয়ে কাজ করি। অর্থাৎ কার পরে কে আসছে সেটা খেয়াল রাখি। ফলে $\{\vec{u}, \vec{v}\}$ এবং $\{\vec{v}, \vec{u}\}$ একই basis হলেও ordered basis হিসেবে আলাদা। তবে অধিকাংশ জায়গাতেই লোকে আর কষ্ট করে ordered basis লেখে না, খালি basis লিখলেই ordered basis বুঝে নেয়।

Exercise 28: Fig 31 আর Fig 32-এ একই vector আঁকা রয়েছে $\vec{w}$. কিন্তু coordinate system-গুলো বিভিন্ন। যেমন ধরো Fig 31-এ যদি আমাদের ordered basis-টা নিই $\{\vec{u}_1, \vec{u}_2\}$ তবে তার সাপেক্ষে $\vec{w}$-কে লিখব লিখব $(3, 2)$. যদি ordered basis-টা নিতাম $\{\vec{u}_2, \vec{u}_1\}$,তবে সেই একই $\vec{w}$-কে লিখতাম $(2, 3)$. এবার বল তো নীচের বিভিন্ন ordered basis-গুলোর জন্য $\vec{w}$-কে কী কী লিখব--

(1) $\{\vec{u}_3, \vec{u}_4\}$, **(2)** $\{\vec{u}_3, \vec{u}_2\}$, **(3)** $\{\vec{u}_4, \vec{u}_1\}$, **(4)** $\{\vec{v}_1, \vec{v}_2\}$, **(5)** $\{\vec{v}_2, \vec{v}_3\}$, **(6)** $\{\vec{v}_4, \vec{v}_3\}$, **(7)** $\{\vec{v}_3, \vec{v}_4\}$. ■

$\mathbb{R}^2$-র বেলায় দেখলাম যে coordinate system আর (ordered) basis একই জিনিস। বুঝতেই পারছ $\mathbb{R}^3$-র বেলাতেও তাই। শুধু তাই নয়, $\mathbb{R}^4$ বা $\mathbb{R}^5$ বা যে কোনো $\mathbb{R}^n$-এর বেলাতেও একই গল্প। যদিও সে সব ক্ষেত্রে গ্রাফ কাগজে ছবি

আঁকার আর প্রশ্ন নেই। এখানেই basis নিয়ে কাজ করার সুবিধা। যেকোনো finite dimensional vector space-ই তাহলে গ্রাফ কাগজ হয়ে দাঁড়ায়, যেকোনো vector-কেই প্রকাশ করা যায় একগুচ্ছ সংখ্যা দিয়ে। এই যে "গুচ্ছ" বললাম একে অংকের ভাষায় বলে **tuple** (টুপ্‌ল্)। যেমন $(3, 4, -5)$ হল একটা 3-tuple. মনে রেখো যে set আর tuple কিন্তু এক জিনিস নয়, set-এর বেলায় কোনটাকে আগে, কোনটাকে পরে লিখছ তাতে কিছু এসে যায় না, tuple-এর বেলায় এসে যায়। যেমন $\{3, 4, -5\}$ আর $\{3, -5, 4\}$ একই set, কিন্তু $(3, 4, -5)$ আর $(3, -5, 4)$ একই 3-tuple নয়।

Example 28: আমরা দেখেছি যে যাবতীয় polynomial, যাদের degree ≤ 2, তারা (এবং zero polynomial-টা) মিলে একটা vector space তৈরী করে, যার নাম দিলাম P_2. তার মানে P_2-র সদস্যরা দেখতে এইরকম--

$$ax^2 + bx + c, \quad (a, b, c \in \mathbb{R}).$$

এর একটা ordered basis দিচ্ছি $\{1 + x, x + x^2, 2x^2\}$. এর সাপেক্ষে $p(x) = 10 + x + 5x^2$-কে একটা 3-tuple আকারে প্রকাশ কর।

SOLUTION: ধরো 3-tuple-টা হবে (a, b, c). মানে

$$10 + x + 4x^2 = a(1 + x) + b(x + x^2) + c(2x^2).$$

বা $10 + x + 5x^2 = a + (a + b)x + (b + 2c)x^2$.

এবার coefficient match করিয়ে পাই $a = 10$, $a + b = 1$ এবং $b + 2c = 5$. তার মানে $a = 10, b = -9$ এবং $c = 7$. সুতরাং উত্তর হল $(10, -9, 7)$. ■

Exercise 29: এখানেও আমরা আগের অংকের P_2 নিয়েই কাজ করব, $p(x)$-ও রাখব আগের মতই। খালি ordered basis-টা নেব $\{x^2 - x, x - 1, 1\}$. এবার 3-tuple-টা কী হবে? ■

Exercise 30: Find the coordinate vector of $\alpha = (1, 3, 1)$ relative to the ordered basis $B = \{(1, 1, 1),\ (1, 1, 0),\ (1, 0, 0)\}$ of $\mathbb{R}^3$.[2] **(2022.CC6.3bii)**

HINT: এখানে **coordinate vector** বলতে বোঝাচ্ছে এমন একটা 3-tuple, ধরো (a, b, c), যার জন্য

$$\alpha = a(1, 1, 1) + b(1, 1, 0) + c(1, 0, 0).$$

■

একটা vector-কে একটা tuple বলে ভাবা অনেক সময়েই বেশ কাজে লাগে। যেমন নীচের অংকটায়।

Example 29: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). Let V be an n-dimensional vector space over a field F with m elements. Then the number of elements in V is

(A) n^m (B) m^m (C) m^n (D) n^n.

(2019.CC6.1g)

SOLUTION: যেহেতু $\dim(V) = n$, তাই একটা basis নিতে পারি--

> let $\{\vec{v}_1, ..., .\vec{v}_n\}$. be a basis of V.

এই basis-এর সাপেক্ষে যে কোনো vector-কেই একটা n-tuple হিসেবে লেখা যায়। বিপরীতপক্ষে, যেকোনো n-tuple থেকেই আমরা একটা vector বানাতে পারি। এইভাবে V এবং যাবতীয় n-tuple-এর set-এর মধ্যে একটা bijection পাচ্ছি--

> Then we have a bijection between all n-tuples of scalars and V:
>
> $$\vec{v} = \alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n \leftrightarrow (\alpha_1, ..., \alpha_n).$$

তাই V-এর সাইজ হবে এরকম যাবতীয় n-tuple-এর সংখ্যা। এই n-tuple-এর মধ্যে n-খানা entry আছে, প্রত্যেকটাকেই যেকোনো scalar থাকতে পারে। একটা entry-তে কী scalar থাকছে, তার সঙ্গে অন্য আরেকটা entry-র কোনো সম্পর্ক নেই। তাই মোট n-tuple-এর সংখ্যা হবে--

> The total number of such tuples is $\underbrace{n \times \cdots \times n}_{m \text{ times}} = m^n$.

■

Exercise 31: Let V be the three-dimensional vector space over the field $\mathbb{Z}_3$. The number of elements of V is

(A) 3 (B) 9 (C) 27 (D) 81

(2022.CC6.1h)
HINT: আগের অংকের মতই। ■

10.3.2 Dimension মানে "degrees of freedom"

Dimension কাকে বলে আমরা শিখেছি। একটা vector space-এর যে কোনো basis-এরই সমান সাইজ হয়, সেই সাইজটাকে বলে vector space-টার dimension. এই বর্ণনা থেকে মনে হয় যেন কোনো vector space-এর dimension বার করতে হলে প্রথমে তার basis বার করে গুণে দেখতে হবে তাতে কটা vector আছে। সেভাবে অবশ্যই dimension বার করা যায়, কিন্তু অনেক সময়েই আরেকভাবে চিন্তা করলে আরো চট্ করে চোখে দেখেই dimension বলে দেওয়া যায়। এই কায়দাটার কোনো প্রচলিত নাম আছে বলে জানি না, কিন্তু physics-এর জগতে degrees of freedom বলে একটা ধারণা আছে, সেটা বস্তুতঃ একই জিনিস, তাই আমরাও কায়দাটার এই নাম দিয়েছি।

Example 30: ধর $V = \{(x, y, z) \ : \ x + y + z = 0\}$. তাহলে $\dim(V)$ কত হবে?

SOLUTION: প্রথমে dimension-এর কথা ভুলে যাও। মনে কর তোমার কাজ হল V-এর কোনো vector-কে যথাসম্ভব সংক্ষেপে প্রকাশ করা। ধরো তোমাকে vector-টার পরিচয় SMS করে পাঠাতে হবে, এবং কতগুলো সংখ্যা পাঠাচ্ছ তার উপর SMS-এর খরচা নির্ভর করছে। যদি x, y, z তিনটেই পাঠাও, তবে অবশ্যই V-এর যেকোনো vector-কেই প্রকাশ করতে পারবে। কিন্তু যেহেতু $x + y + z = 0$ বলা আছে, তাই খালি x, y পাঠানোই যথেষ্ট। x, y পেলেই তোমার বন্ধু তা থেকে $z = -(x + y)$ বার করে নিতে পারবে। লক্ষ্য কর এর চেয়ে কম সংখ্যা পাঠিয়ে তুমি কাজ সারতে পারবে না। তাই dimension হবে 2.
তুমি অবশ্যই x, y না পাঠিয়ে y, z-ও পাঠাতে পারতে, সেক্ষেত্রে বন্ধু $x = -(y + z)$ বার করে নিত। কী পাঠাবে সেটা বন্ধুর সঙ্গে আগেভাগেই ঠিক করে রাখতে হবে। চাইলে তুমি $x + y, x + z$-ও পাঠাতে পারো, তা থেকেও x, y, z সবই উদ্ধার করা যাবে। কিন্তু যেভাবেই যা কর না কেন, সব সময়েই তোমাকে দুটো সংখ্যাই পাঠাতে হবে। এবং এই 2 সংখ্যাটাই হবে $\dim(V)$. ■

এখানে x, y, z এই তিনটে সংখ্যার মধ্যে সবাই স্বাধীন ছিল না, যেমন $x = 2$ আর $y = 3$ নিলে z যা খুশি কিছু হতে পারে না, ওকে $-(x + y) = -5$ হতেই হবে। তাই এখানে যেন x, y হল free কিন্তু z মোটেই free নয়। অবশ্য তুমি একইভাবে y, z-কে free ভাবতে পারো, সেক্ষেত্রে x আর free থাকবে না। কিন্তু সবসময়েই free-দের মোট সংখ্যা একই থাকবে, এবং সেটাই হবে dimension. সেই থেকেই degrees of freedom কথাটার উদ্ভব।

Exercise 32: Degrees of freedom-এর কায়দাটা লাগিয়ে চট্ করে বল তো নীচের vector space-গুলোর dimension কী কী হবে! প্রতি ক্ষেত্রেই field-টা নেব $\mathbb{R}$. তার মানে তোমাকে ভাবতে হবে সবচেয়ে কম কতগুলো real number দিয়ে তুমি যেকোনো vector-কে প্রকাশ করতে পারছ।

1. $V_1 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \ : \ x_1 = x_2, \ x_2 = x_3, \ x_3 = x_1\}$.
2. $V_2 = \{(z_1, z_2) \in \mathbb{C}^2 \ : \ Re(z_1) = Im(z_2)\}$.

■

এবার একটা বিশ্রীরকমের ঠকানো প্রশ্ন আসছে, সুতরাং সাবধান!

Exercise 33: আগের অংকের V_2-টা দ্যাখো, কিন্তু এবার যদি field-টা $\mathbb{C}$ নিয়ে dimension বার করতে বলি তবে?
■

Exercise 34: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). Let V be the real vector space of all 3×3 real matrices, and W be the subspace of V consisting of all symmetric matrices. Then the dimension of W is

(A) 9 (B) 6 (C) 3 (D) 8

(2022.CC6.1g)
HINT: এখানে matrix-গুলো 3×3 সাইজের। যদি ওদের entry-গুলোকে একেকটা অক্ষর দিয়ে লিখি, তবে মোট 9-টা অক্ষর লাগবে। যদি symmetric matrix হয়, তবে কটা অক্ষর লাগবে? সেই সংখ্যাটাই হল উত্তর। সংক্ষেপে প্রমাণ করার সময়ে প্রতিটা অক্ষরের অবদানকে আলাদা করে নিলেই একটা basis পেয়ে যাবে। যেমন যদি 2×2 হত, তবে symmetric matrix বোঝাতে তিনটে অক্ষরই যথেষ্ট, $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$. তাই dimension হত 3. অবদানগুলোকে আলাদা করলে হয়

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

এ থেকেই পাবে যে

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

হল একটা generating set. এটা যে একটা linearly independent set-ও বটে, সেটা দেখার জন্য যা খুশি a, b, c নিয়ে linear combination বানাও, তবে সেই $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$-ই ফিরে পাবে, সেটা যদি $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ হয়, তবে যে $a = b = c = 0$ হতে বাধ্য, সে তো দেখাই যাচ্ছে! তার মানে ওই তিনটে matrix-এর set-টা হল একটা basis. তাই dimension হচ্ছে 3. এই উদাহরণটা মাথায় রেখে এবার আমাদের এই 3×3 অংকটা করো। ■

Exercise 35: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). For the subspaces $U = \left\{ \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \right\}$ and $V = \left\{ \begin{bmatrix} a & b \\ c & -c \end{bmatrix} \right\}$ the dim U and dim V are

(A) (2,1) (B) (2,3) (C) (2,2) (D) (1,3).

(2021.CC6.1g)

HINT: এখানে U হল যাবতীয় $\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}$ চেহারা matrix, যেখানে c, d হল যা খুশি scalar. একইভাবে V হল যাবতীয় $\begin{bmatrix} a & b \\ c & -c \end{bmatrix}$ চেহারার matrix, যেখানে a, b, c হল যা খুশি scalar. এবার option-গুলোর মানে বলি-- প্রথম option-টা বলছে $\dim(U) = 2$ এবং $\dim(V) = 1$, এইরকম। ■

DAY 11
Dimension theorem

আমরা আগেই বলেছি যে একটা vector space-এর একাধিক basis থাকতেই পারে, কিন্তু সবগুলো basis-এরই সাইজ একই হয়। তাই আমরা বলতে পারি যে এই সাইজটা basis-গুলোর নিজস্ব বৈশিষ্ট্য নয়, ওটা vector space-টার নিজস্ব বৈশিষ্ট্য। এই সাইজটাকে যে vector space-টার dimension বলে সে কথাও বলেছি। এখনো অবশ্য আমরা প্রমাণ করি নি যে, একটা vector space-এর সব basis-এরই সাইজ সমান। এইটা একটা অত্যন্ত গুরুত্বপূর্ণ theorem, যার নাম dimension theorem. এবার সেটাই আমরা প্রমাণ করব। এর জন্য replacement বলে একটা ধারণা আগে বুঝে রাখা দরকার।

11.1 Replacement

ব্যাপারটা বোঝার জন্য একটা খেলার উপমা মাথায় রাখলে সুবিধা হবে। খেলাটা এইরকম--তোমাকে দুটো ইংরাজি শব্দ দেওয়া হবে, যাদের অক্ষরসংখ্যা সমান। যেমন CAT আর DOG. তোমার কাজ হল ধাপে ধাপে একটা শব্দকে অন্যটায় পরিণত করা। প্রতি ধাপে তুমি খালি একটা করে অক্ষরকে পালটাতে পারবে, এবং খেয়াল রাখতে হবে যে পালটানোর পর নতুন যে শব্দটা পাবে সেটা যেন অর্থপূর্ণ হয়। একটা উদাহরণ দেখিয়েছি Fig 33-এ। কাজটা Fig 34-এর মত করেও করা যেত।

একইভাবে এবার আমরা ধাপে ধাপে একটা basis থেকে আরেকটা basis-এ যাব। প্রত্যেকটা ধাপে একটা করে vector-কে replace করব (মানে একটা vector-এর জায়গায় আরেকটাকে আনব)। কাজটা এমনভাবে করতে হবে, যাতে প্রতি ধাপেই আমরা একটা করে basis পাই। একটু খুঁটিয়ে দেখা যাক।

Example 31: Fig 35-এ চারটে vector রয়েছে। এদের মধ্যে $\{\vec{u}_1, \vec{u}_2\}$ একটা basis. এরা যেন দুজন পাহারাদার। এদের মধ্যে $\vec{u}_1$ ছুটিতে বাড়ি যেতে চায়। তার বদলি হিসেবে দুজন দরখাস্ত দিয়েছে, $\vec{v}_1$ আর $\vec{v}_2$. এদের মধ্যে কাকে $\vec{u}_1$-এর জায়গায় নিয়োগ করা চলবে?

SOLUTION: $\vec{v}_1$-কে দিয়ে হবে না, কারণ $\vec{v}_1$ আর $\vec{u}_2$ একই লাইন বরাবর, ফলে ওরা দুজনে মিলে হাজার চেষ্টা করেও ওই লাইনের বাইরে কোনো কিছুর তদারক করতে পারবে না। এতক্ষণ এই লাইনের বাইরে তাকানোর দায়িত্ব ছিল $\vec{u}_1$-এর। তার জায়গায় যে আসবে তার মধ্যেও $\vec{u}_1$-এই গুণটা থাকতে হবে। সে যদি পুরোই $\vec{u}_2$-এর লাইনেই থাকে, তবে আর তাকে দিয়ে বাড়তি কী কাজ হবে? সুতরাং এখানে $\vec{u}_1$-এর বদলি হিসেবে $\vec{v}_2$-কে নিয়োগ করা যাবে।

$\vec{v}_2$-র প্রধান যে যোগ্যতার জন্য তাকে $\vec{u}_1$-এর স্থলাভিষিক্ত করা গেল, সেটা এই যে $\vec{v}_2$-টা $\vec{u}_2$-র span-এর মধ্যে নেই। তার মানে যদি $\vec{v}_2$-কে আমাদের মূল basis দিয়ে ভেঙে লিখি এইভাবে--

$$\vec{v}_2 = a_1\vec{u}_1 + a_2\vec{u}_2,$$

Fig 33

C A T

C O T

D O T

D O G

Fig 34

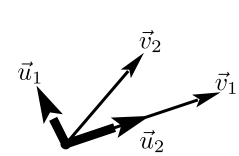
C A T

C O T

C O G

D O G

Fig 35



তবে $a_1 \neq 0$. অর্থাৎ যার বদলে ও আসছে তার খানিকটা অংশ ওর মধ্যে আছে। ■

এই উদাহরণটা যদি বুঝে থাকো তবে নীচের অংকটা সহজেই হওয়া উচিত।

Example 32: Suppose $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ is a basis of a vector space V over a field F and a nonzero vector β of V is expressed as

$$\beta = c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 + c_4\alpha_4; \quad c_i \in F(i = 1, 2, 3, 4),$$

then if $c_2 \neq 0$, prove that $\{\alpha_1, \beta, \alpha_3, \alpha_4\}$ is a new basis.[5] **(2005.13a)**

SOLUTION: প্রথমে ছিল চারজন পাহারাদার--

> Let $S = \{\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3, \vec{\alpha}_4\}$.

প্রশ্নে না থাকলেও, আমরা vector-দের মাথায় তীর চিহ্ন বসিয়ে নিয়েছি (ওগুলো আমাদের পাহারাদারদের উর্দির অংশ!)। এবার এদের মধ্যে $\vec{\alpha}_2$ গেল ছুটিতে, তার জায়গায় এল $\vec{\beta}$.

> Let $T = \{\vec{\alpha}_1, \vec{\beta}, \vec{\alpha}_3, \vec{\alpha}_4\}$.

এবার দেখাতে হবে যে T একটা basis, মানে একাধারে linearly independent এবং generating set.

> Step 1: Shall show that T is a linearly independent set.

আমরা linear independence দেখানোর প্রচলিত কায়দা লাগাব--

> Let $a_1, b, a_3, a_4 \in F$ be such that
>
> $$a_1\vec{\alpha}_1 + b\vec{\beta} + a_3\vec{\alpha}_3 + a_4\vec{\alpha}_4 = \vec{0}.$$

দেখাতে হবে যে $a_1 = b = a_3 = a_4 = 0$. এবার আমরা $\vec{\beta}$-কে বিদায় করার ব্যবস্থা করি--

> Putting $\vec{\beta} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + c_3\vec{\alpha}_3 + c_4\vec{\alpha}_4$, we have
>
> $$a_1\vec{\alpha}_1 + b(c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + c_3\vec{\alpha}_3 + c_4\vec{\alpha}_4) + a_3\vec{\alpha}_3 + a_4\vec{\alpha}_4 = 0,$$
>
> or
>
> $$(a_1 + bc_1)\vec{\alpha}_1 + bc_2\vec{\alpha}_2 + (a_3 + bc_3)\vec{\alpha}_3 + (a_4 + bc_4)\vec{\alpha}_4 = 0.$$

কিন্তু $\vec{\alpha}_i$-রা তো linearly independent, অতএব--

> $\because S$ is linearly independent,
>
> $$\therefore a_1 + bc_1 = bc_2 = a_3 + bc_3 = a_4 + bc_4 = 0.$$
>
> But $c_2 \neq 0$. So $b = 0$. Hence $a_1 = a_3 = a_4 = 0$.
> Thus we have $a_1 = b = a_3 = a_4 = 0$, proving linear independence of T.

∎

এবার পরের ধাপ।

Step 2: Shall show that T is a generating set, ie,

$$\forall \vec{v} \in V \ \ \vec{v} \in \text{span}(T).$$

Take any $\vec{v} \in V$.
$\because S$ is a generating set,

$$\therefore \vec{v} = d_1\vec{\alpha}_1 + d_2\vec{\alpha}_2 + d_3\vec{\alpha}_3 + d_4\vec{\alpha}_4$$

for some $d_1, ..., d_4 \in F$.
Now, $\because c_2 \neq 0$,

$$\therefore \vec{\alpha}_2 = \frac{1}{c_2}(\vec{\beta} - c_1\vec{\alpha}_1 - c_3\vec{\alpha}_3 - c_4\vec{\alpha}_4).$$

So

$$\begin{aligned} \vec{v} &= d_1\vec{\alpha}_1 + d_2\vec{\alpha}_2 + d_3\vec{\alpha}_3 + d_4\vec{\alpha}_4 \end{aligned}$$

এইবার আমরা $\vec{\alpha}_2$-এর জায়গায় $\vec{\beta}$-কে আনার চেষ্টা করব।

$$\begin{aligned} &= d_1\vec{\alpha}_1 + \frac{d_2}{c_2}(\vec{\beta} - c_1\vec{\alpha}_1 - c_3\vec{\alpha}_3 - c_4\vec{\alpha}_4) + d_3\vec{\alpha}_3 + d_4\vec{\alpha}_4 \\ &\in \text{span}(T). \end{aligned}$$

Combining the two steps we see that T is a basis of V, as required.

■

এর বিপরীত দিকটাও কিন্তু সত্যি। সেটা দেখালাম নীচের অংকটায়। এই প্রমাণটা অনেক বেশী সহজ।

Example 33: Suppose $\{\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3, \vec{\alpha}_4\}$ is a basis of a vector space V over a field F and a vector $\vec{\beta}$ of V is expressed as

$$\vec{\beta} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + c_3\vec{\alpha}_3 + c_4\vec{\alpha}_4; \quad c_i \in F(i = 1, 2, 3, 4).$$

If $\{\vec{\alpha}_1, \vec{\beta}, \vec{\alpha}_3, \vec{\alpha}_4\}$ is also a basis, then prove that $c_2 \neq 0$.
SOLUTION:

Let, if possible, $c_2 = 0$.
Then

$$\vec{\beta} = c_1\vec{\alpha}_1 + c_3\vec{\alpha}_3 + c_4\vec{\alpha}_4 \in \text{span}\{\vec{\alpha}_1, \vec{\alpha}_3, \vec{\alpha}_4\}.$$

So $\{\vec{\alpha}_1, \vec{\beta}, \vec{\alpha}_3, \vec{\alpha}_4\}$ is not linearly independent, and hence cannot be a basis ($\Rightarrow\Leftarrow$).

ওই $\Rightarrow\Leftarrow$ চিহ্নটার মানে হল contradiction. ■

এই দুটো অংক থেকে যা শেখা গেল, তাদেরকেই একত্রে বলে replacement theorem–

Replacement theorem

Let V be any vector space, and let $\{\vec{u}_1, ..., \vec{u}_n\}$ be any given basis of V. Let $\vec{v} \in V$ be some vector and let $i \in \{1, ..., n\}$. Suppose that

$$\vec{v} = a_1\vec{u}_1 + \cdots + a_n\vec{u}_n.$$

Then

$$\{\vec{u}_1, ..., \vec{u}_{i-1}, \vec{v}, \vec{u}_{i+1}, ..., \vec{u}_n\}$$

is a basis of V if and only $a_i \neq 0$.

Example 34: Prove that the set $S = \{(1,1,0),\ (1,0,1),\ (0,1,1)\}$ is a basis of the vector space $\mathbb{R}^3$. Show that the vector $(1,1,1)$ may replace any one of the vectors of the set S to form a new basis.[2+3] **(2020.CC6.3e)**

SOLUTION: প্রথমে আমরা দেখাব যে S হল একটা spanning set.

Let $(x, y, z) \in \mathbb{R}^3$. We want $a, b, c \in \mathbb{R}$ such that

$$a(1,1,0) + b(1,0,1) + c(0,1,1) = (x,y,z).$$

Then

$$\begin{aligned} a + b &= x, \\ a + c &= y, \\ b + c &= z. \end{aligned}$$

Then $a + b + c = \frac{x+y+z}{2}$ and so $a = (a+b+c) - (b+c) = \frac{x+y+z}{2} - z = \frac{x+y-z}{2}$. Similarly, $b = \frac{x-y+z}{2}$ and $c = \frac{-x+y+z}{2}$.

যেহেতু যেকোনো (x, y, z)-এর জন্যই a, b, c পাওয়া যাচ্ছে, তাই--

So S spans $\mathbb{R}^3$.

এবার linearly independent দেখাব, কিন্তু তাতে আর নতুন করে খাটতে হবে না।

Also, if we take $(x, y, z) = (0, 0, 0)$, then $a = \frac{0+0-0}{2} = 0$. Similarly, $b = c = 0$.

Hence S is also linearly independent.

Thus, S is a basis of $\mathbb{R}^3$, as required.

এবার replacement theorem লাগানোর পালা। সেখানেও নতুন করে বেশী খাটতে হবে না--

> Taking $(x, y, z) = (1, 1, 1)$, we have $a = \frac{1+1-1}{2} = \frac{1}{2}$. Similarly, $b = c = \frac{1}{2}$.
> Since in the expansion
>
> $$(1, 1, 1) = \frac{1}{2}(1, 1, 0) + \frac{1}{2}(1, 0, 1) + \frac{1}{2}(0, 1, 1)$$
>
> all the coefficients are nonzero, hence by replacement theorem, $(1, 1, 1)$ may replace each of the basis vectors.

অবশ্য এই অংকটা এতই সহজ, এখানে replacement theorem ছাড়াই করে দেওয়া যায়। সেরকম একটা কায়দা দেখাচ্ছি।

বিকল্প পদ্ধতি

> Let $\vec{v}_1 = (1, 1, 0)$, $\vec{v}_2 = (1, 0, 1)$ and $\vec{v}_3 = (0, 1, 1)$.
> Also let $\vec{v} = (1, 1, 1)$.
> Then $\vec{v}_1 + \vec{v}_2 + \vec{v}_3 = 2\vec{v}$.
> So $\vec{v}_1 = 2\vec{v} - \vec{v}_2 - \vec{v}_3 \in \operatorname{span}\{\vec{v}, \vec{v}_2, \vec{v}_3\}$.
> Hence $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \subseteq \operatorname{span}\{\vec{v}, \vec{v}_2, \vec{v}_3\}$.
> But $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \mathbb{R}^3$.
> Hence $\operatorname{span}\{\vec{v}, \vec{v}_2, \vec{v}_3\} = \mathbb{R}^3$.

এবার linearly independent দেখাব--

> Also $\{\vec{v}, \vec{v}_2, \vec{v}_3\}$ is linearly independent.
>
> ⟦ Because:
> If $a\vec{v} + b\vec{v}_2 + c\vec{v}_3 = \vec{0}$, then
> $(a + b, a + c, a + b + c) = (0, 0, 0)$, and hence $a = b = c = 0$.
> ⟧
>
> So $\{\vec{v}, \vec{v}_2, \vec{v}_3\}$ is a basis of $\mathbb{R}^3$. Thus $\vec{v}$ can replace $\vec{v}_1$.
> Similarly, $\vec{v}$ can replace $\vec{v}_2$ and $\vec{v}_3$ also, as required.

■

Exercise 36: ফের আগের অংকটার দিকে তাকাও। এবার বলো তো $(1, 3, 2)$ দিয়ে কি $(1, 0, 1)$-কে replace করা যাবে? ■

Exercise 37: Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be a basis of a vector space, V. Let $\vec{v} \in V$. Then is the following statement true?

"$\vec{v}_n$ may be replaced by $\vec{v}$ if and only if $\vec{v} \notin \operatorname{span}\{\vec{v}_1, ..., \vec{v}_{n-1}\}$."

■

আগেই বলেছি যে, কোনো vector space-এর যে কোনো দুটো basis-এর সাইজ সমান হতে বাধ্য। এর প্রমাণটা হয় replacement theorem দিয়ে। এবার সেটা বোঝা যাক।

11.2 Dimension

আমরা প্রথমে যেকোনো দুটো basis নিয়ে শুরু করি, B আর C. দেখাতে হবে যে B আর C দুজনেরই সাইজ সমান হতে বাধ্য। প্রমাণের খুঁটিনাটি পরে বলছি। আগে প্রমাণের আদলটা বলে নিই। আমরা contradiction-এর পথে এগোব। ধরে নেব যে B আর C-এর সাইজ অসমান, ধরো B-টা ছোটো আর C-টা বড়। সেখান থেকে কোনো একটা contradiction-এ পৌঁছব। বোঝার সুবিধার জন্য একটা উদাহরণ নিই--

$$B = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$$

আর

$$C = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}.$$

এবার আমরা ধাপে ধাপে replacement theorem লাগাব। C-র কোনো একটা vector দিয়ে B-র কোনো একটা vector-কে replace করে দেব। ধরো $\vec{v}_1$ দিয়ে $\vec{u}_1$-কে replace করে দিলাম। এর ফলে B-টা বদলে হয়ে গেল

$$B_1 = \{\vec{v}_1, \vec{u}_2, \vec{u}_3\}.$$

এরপর একইভাবে আরেকটা vector ঢোকাব C থেকে B_1-এর মধ্যে। ধরো এবার $\vec{u}_2$-কে replace করে দিলাম $\vec{v}_2$ দিয়ে। ফলে পেলাম

$$B_2 = \{\vec{v}_1, \vec{v}_2, \vec{u}_3\}.$$

লক্ষ কর কীভাবে $\vec{u}$-দের একে একে ঘাড়ে ধাক্কা দিয়ে বিদায় করে দিচ্ছি। ঠিক যেমন ভারতের স্থানীয় ব্যবসায়ীদের দোকানগুলোকে বহুজাতিক সংস্থারা একে একে হটিয়ে দেয়। এখনও B_2-এর মধ্যে $\vec{u}_3$ আছে, যদি সেটাকেও $\vec{v}_3$ দিয়ে replace করে দিই, তবে $\vec{u}$-রা সম্পূর্ণ নিশ্চিহ্ন হয়ে গিয়ে পড়ে থাকবে খালি

$$B_3 = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}.$$

প্রতিধাপেই আমরা replacement theorem লাগিয়েছি। তাই প্রতিটা B_i-ই একেকটা basis. তার মানে B_3-ও একটা basis. লক্ষ্য করো C-এর মধ্যে যে সব vector এখনও ব্যবহার হয়নি, $\vec{v}_4$ পড়ে আছে। যেহেতু B_3 একটা basis তাই অবশ্যই $\vec{v}_4 \in \text{span}(B_3)$ হতে বাধ্য। কিন্তু তার মানে তো

$$\vec{v}_4 \in \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\},$$

যেটা অসম্ভব, কারণ C তো একটা linearly independent set, তার কোনো সদস্যই বাকিদের পয়সায় বসে খেতে পারে না!

এই হল প্রমাণের আদলটা। অবশ্য এর মধ্যে একটা ফাঁক থেকে গেছে। Replacement theorem তো আর যেখানে সেখানে লাগিয়ে দেওয়া যায় না, তার তো একটা শর্ত আছে। এই যে আমরা গড়গড় করে বলে গেলাম $\vec{v}_1$ দিয়ে $\vec{u}_1$-কে replace কর, $\vec{v}_2$ দিয়ে $\vec{u}_2$-কে replace করো ইত্যাদি, সেই প্রতিটা ধাপে replacement theorem-এর শর্তটা পরীক্ষা করে দেখাতে হবে। তবে কাজটা মোটেই শক্ত নয়, একটা ধাপে দেখালেই আসলে চলবে, কারণ ধাপগুলো সবই একই রকম। সব মিলিয়ে কী পেলাম সেটা গুছিয়ে একটা theorem আকারে লিখে ফেলি।

Dimension theorem

Let V be a vector space with at least one finite basis. Then every basis must be finite and of the same size.

Proof:

> We are given that a finite basis exists.
> Let us take any other basis.
> We call them B and C so that $|B| \leq |C|$.

এইখানটা একটু বুঝে নেওয়া যাক। আমরা দুটো basis নিয়ে কাজ করছি, তার মধ্যে একটা বলা আছে finite. অন্যটা আমাদের নেওয়া, সেটা finite নাকি infinite সে বিষয়ে এখনও কিছু জানি না। আমরা এদের মধ্যে ছোটোটার নাম দিলাম B আর অন্যটার নাম দিলাম C. সেই জন্য $|B| \leq |C|$ লিখেছি। আমরা দেখাব $|C| = |B|$ হবে।

> Then B is a finite basis. Let $n = |B|$.
> Shall show $|C| = n$.
> Let $B = \{\vec{u}_1, ..., \vec{u}_n\}$.
> We pick any $\vec{v}_1 \in C$.

এইবার আমরা $\vec{v}_1$-কে B-র মধ্যে ঢোকাব, অর্থাৎ কোনো একটা $\vec{u}_i$-কে $\vec{v}_1$ দিয়ে replace করব। তার জন্য replacement theorem-এর শর্তটা পরীক্ষা করে দেখতে হবে। প্রথম কাজ হল $\vec{v}_1$-কে B-এর vector-গুলো দিয়ে প্রকাশ করা--

> Then $\vec{v}_1$ can be written as
>
> $$\vec{v}_1 = a_1\vec{u}_1 + \cdots + a_n\vec{u}_n.$$
>
> for some scalars $a_1, ..., a_n$.

এবার দেখতে হবে এই scalar-গুলোর মধ্যে nonzero কিছু পাওয়া যায় কিনা--

> $\because C$ is a basis,
> $\therefore \vec{v}_1 \neq \vec{0}$. So at least one a_i must be nonzero.
> W.l.g., let $a_1 \neq 0$.

এখানে w.l.g. অর্থাৎ without loss of generality লিখেছি, কারণ $a_1 \neq 0$ ধরে নিয়েছি বটে, কিন্তু যদি অন্য কোনো $a_i \neq 0$ হত, তাহলেও যুক্তিটা একেবারেই একইরকম থাকত (যদিও লেখাটা একটু বেশী জটিল হত)।
এইবার replacement theorem লাগিয়ে পেলাম আমাদের নতুন basis–

> By replacement theorem, $B_1 = \{\vec{v}_1, \vec{u}_2, ..., \vec{u}_n\}$ is a basis.

একটা $\vec{v}$-কে ঢুকিয়েছি $\vec{u}$-দের দঙ্গলের মধ্যে। এবার আরেকটাকে ঢোকাব। যুক্তিটা প্রায় একই রকম, খালি একটা নতুন প্যাঁচ আছে, তাই এখনই "similarly" লিখে ছেড়ে দিতে পারছি না।
প্রথমে আরেকটা $\vec{v}$ নিই--

> Let $\vec{v}_2 \in C \quad (\vec{v}_2 \neq \vec{v}_1)$.

আবার replacement theorem লাগানোর জন্য এই $\vec{v}_2$-কে B_1-এর vector-গুলো দিয়ে প্রকাশ করি--

> Then $\vec{v}_2 = b_1\vec{v}_1 + b_2\vec{u}_2 + \cdots + b_n\vec{u}_n$ for some scalars $b_1, ..., b_n$.

এবার আগের ধাপের মতই দেখাতে পারি যে এই $b_1, ..., b_n$-এর মধ্যে অন্ততঃ একটা nonzero আছে। কিন্তু সেটুকু করলেই খালি হবে না। কারণ $b_1 \neq 0$ পেয়ে আমাদের কোনো লাভ নেই, কারণ সেক্ষেত্রে $\vec{v}_2$ দিয়ে $\vec{v}_1$-টা replaced হয়ে যাবে! তাতে করে $\vec{u}$-দের সংখ্যা কমবে না। তাই আমাদের দেখাতে হবে যে $b_2, ..., b_n$-দের মধ্যে অন্ততঃ একজন nonzero থাকতে বাধ্য। এইটুকুই খালি নতুন প্যাঁচ--

> $\because C$ is a linearly independent set,
> $\therefore \vec{v}_2 \notin \text{span}\{\vec{v}_1\}$.
> $\therefore \exists i \in \{2, ..., n\} \quad b_i \neq 0$.
> W.l.g., $i = 2$.
> By replacement theorem, $B_2 = \{\vec{v}_1, \vec{v}_2, \vec{u}_3, ..., \vec{u}_n\}$ is a basis.

এইবার similarly লেখা চলবে--

> Continuing similarly for n steps we get a basis $B_n = \{\vec{v}_1, ..., \vec{v}_n\} \subseteq C$.
> Then $C = B_n$.
>
> [[Because:
> Clearly, $B_n \subseteq C$.
> Let, if possible, $\exists \vec{v} \in C \setminus B_n$.
> $\because B_n$ is a basis, $\therefore \vec{v} \in \text{span}(B_n)$.
> Thus $\vec{v} \in C$ belongs to the span of some other vectors in $C (\Rightarrow\Leftarrow \because C$ is linearly independent).
>]]
>
> Thus $|C| = |B_n| = n$, as required.
> Since C was any arbitrary basis, hence this proves that all bases must have the same size.

[Q.E.D]

Exercise 38: $\{(1, 2, 3), (2, 3, 4)\}$ কি $\mathbb{R}^3$-র একটা basis? এক সেকেণ্ডের মধ্যে উত্তর দাও। ■

DAY 12
Extension আর contraction

একটা রূপকথার গল্পে দুই পণ্ডিতের ঝগড়ার কথা পড়েছিলাম। ঝগড়া বেঁধেছিল শুয়োর নামক প্রাণীটির উৎপত্তি কী করে হয়েছে তাই নিয়ে। এক পণ্ডিতের মতে ইঁদুর বড় বড় হতে হতে শুয়োর হয়েছে, আরেক পণ্ডিতের ধারণা হাতি ছোটো হতে হতে শুয়োর হয়েছে। আমাদের basis-এর অবস্থাটাও কতকটা সেইরকম (Fig 36)। একটা basis কি কোনো linearly independent set বাড়তে বাড়তে সৃষ্টি হয়? নাকি generating set ছোটো হতে হতে? শুয়োরের বেলায় অবশ্য দুই পণ্ডিতের ধারণাই ভুল

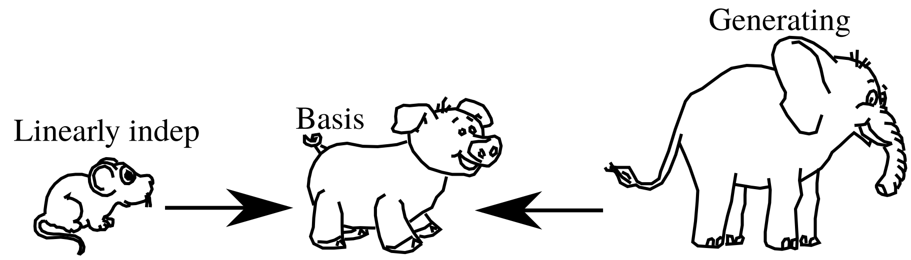


Fig 36

ছিল। কিন্তু basis-এর বেলায় দুটো ধারণাই সঠিক। এবার আমরা সেটাই দেখব। প্রথমে দেখি একটা linearly independent set বড় হতে হতে কীভাবে basis তৈরী হয়।

12.1 Extension theorem

Fig 37-এর ছবিটা আমরা আগেই দেখেছি। এই কাল্পনিক চিত্রটাকে মাথায় রাখলে basis বার করা সহজ হয়। একটা উদাহরণ দ্যাখো।

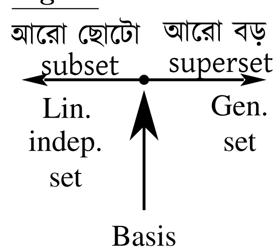
Example 35: ধরো $\mathbb{R}^3$-র basis বার করতে বলল। তবে গোড়ায় যেকোনো একটা nonzero vector নিয়ে শুরু কর, যেমন $(1,0,0)$. তাহলে একটা linearly independent set পেলাম $\{(1,0,0)\}$. এটা কি একটা generating set-ও বটে? না, কারণ এই একজন পাহারাদার তো খালি তার নিজের লাইন বরাবর পাহারা দিতে পারে। কিন্তু $\mathbb{R}^3$ তো খালি একটা লাইন মাত্র নয়! বেশ কথা, তাহলে এই লাইনের বাইরে কিছু একটা vector নাও, যেমন $(0,1,0)$. সুতরাং আমাদের linearly independent set-টা আরেকটু বেড়ে হল

$$\{(1,0,0),\ (0,1,0)\}.$$

এটা কি একটা generating set? উঁহু, এরা দুজনে মিলে পুরো xy-plane-কে পাহারা দিতে পারে, কিন্তু $\mathbb{R}^3$ তো খালি একটা plane মাত্র নয়! অতএব এই plane-এর বাইরে থেকে কিছু একটা vector নাও, ধরো $(0,0,1)$. এইবার একটা basis হল! ■

খানিকক্ষণ চিন্তা করলেই বুঝবে যে, এই কায়দায় যেকোনো linearly independent set-কেই বাড়াতে বাড়াতে basis বানিয়ে ফেলা যায়। যতক্ষণ না basis হয়, ততক্ষণ তার span-এর বাইরে থেকে যা খুশি একটা করে vector-কে set-এর মধ্যে ঢুকিয়ে যেতে হবে। যেহেতু finite dimensional vector space নিয়ে কাজ করছি, অতএব কোনো এক সময়ে এই কাজটা শেষ হতে বাধ্য।
এই ধারণাটাকে গুছিয়ে লিখলেই আমরা পাই নীচের theorem-টা।

Fig 37



Extension theorem

If V is any finite dimensional vector space, and $S \subseteq V$ is any linearly independent set, then there must exist a basis $B \supseteq S$.

প্রমাণ রয়েছে নীচের অংকটার সমাধানে।

Example 36: Prove that a set of linearly independent vectors of a finite dimensional vector space is either a basis or can be extended to a basis of the vector space.[5] **(2011.11a,2008,2007,2004)**

SOLUTION:

Let V be a finite dimensional vector space.

Let $S \subseteq V$ be a linearly independent set.

We fix some basis B of V.

Since V finite dimensional, hence B is a finite set.

Case 1: $B \subseteq \text{span}(S)$:

Then $V = \text{span}(B) \subseteq \text{span}(S)$.

Hence S is also a generating set, and hence a basis of V.

Case 2: $B \not\subseteq \text{span}(S)$:

Then we can find some $\vec{b} \in B$ which is not in $\text{span}(S)$.

Then $S_1 = S \cup \{\vec{b}\}$ is a linearly independent set.

⟦ Because:

If $p_1, ..., p_k, q \in F$ are such that

$$\sum_j p_j \vec{s}_j + q\vec{b} = \vec{0},$$

then $q = 0$, because otherwise

$$\vec{b} = \sum_j \frac{p_j}{q} \vec{s}_j \in \text{span}(S)(\Rightarrow\Leftarrow).$$

সুতরাং $\vec{b}$-এর coefficient-টাকে বিদায় করা গেল। বাকিগুলো নিয়ে সমস্যা নেই, কারণ S তো linearly independent ছিলই--

So

$$\sum_j p_j \vec{s}_j = \vec{0},$$

and hence, by linear independence of S,

$$p_1 = \cdots = p_k = 0.$$

]]

If S_1 is a basis, then we are done. Otherwise, we can again find a basis vector outside its span, and grow S_1 further. This process must stop after finitely many steps, since B is finite.

■

12.1.1 ফলশ্রুতি

Extension theorem জিনিসটা খুবই দরকারী। মুরগীর ডিমে তা দিয়ে যেমন মুরগী[2] হয়, তেমনি linearly independent set দেখলেই মনে মনে ভাববে--এটা যেন basis-এর ডিম, দরকার বুঝলেই এটাকে extend করে একটা basis বানিয়ে ফেলব।

Extension theorem-এর তিনটে ফলশ্রুতি তো না বললেই নয়। প্রথমটার উল্লেখ আগেই করেছিলাম। একটা linearly independent set-কে যখন আর বাড়ানো যায় না (মানে আর একটাও নতুন vector ঢোকালে set-টা আর linearly independent থাকে না), তখন বুঝতে হবে যে set-টা আসলে একটা basis. যে linearly independent set-কে আর বাড়ানো যায় না, তার আরেক নাম হল **maximal linearly independent set**. সুতরাং--

THEOREM

A subset of a finite dimensional vector space is a basis if and only if it is a maximal linearly independent set.

প্রমাণ তো বুঝতেই পারছ--যদি basis না হত, তবে তো বাড়ানোই যেত, সেটাই আগের অংকে দেখালাম। আর যদি basis হয়, তবে সকল vector-ই তো ওর span-এর মধ্যে, সুতরাং ওদের কাউকে ঢোকালে সেই set-টা আর linearly independent থাকে কী করে?

নীচের theorem-টাও একই জিনিস, খালি dimension ব্যবহার করে লেখা--

THEOREM

Let V be a vector space of dimension n. Let $S \subseteq V$ be a linearly independent set. Then

1. $|S| \leq n$.
2. $|S| = n \iff S$ is a basis of V.

যদি S একটা linearly independent set হয়, তবে ওকে extend করে একটা basis বানানো যাবে। ধরো এই basis-টার নাম দিলাম B. তাহলে $S \subseteq B$, কারণ S-কে extend করেই তো B-কে পেলাম। এদিকে B যেহেতু basis, সুতরাং ওর সাইজ হবে V-এর dimension, মানে n-এর সমান। সুতরাং $|S| \leq |B| \leq n$.

[2] মোরগও হতে পারে অবশ্য!

যদি S একটা basis হয়, তবে তো ওর সাইজ n হতেই হবে, কারণ $\dim(V) = n$. আর বিপরীতপক্ষে যদি $|S| = n$ হয়, তার মানে S-কে আর extend করার জায়গা নেই। তার মানে S আসলে একটা basis হয়েই আছে। এই শেষ অংশের যুক্তিটাকেই বিশদ করে লিখতে বলেছে নীচের অংকে।

Example 37: Prove that any linearly independent set of n vectors of a vector space V over $\mathbb{R}$ is a basis of V if $dim\ V = n$. [4] **(2021, 2012)**

SOLUTION:

> Let V be a vector space of dimension n.
> Let S be a linear independent set of size n.
> Shall show that S must be a basis. Let, if possible, S be not a basis.

Basis মানে linearly independent আর generating. তার মধ্যে S তো linearly independent আছেই। সুতরাং S যদি basis না হয়, তবে বুঝতে হবে যে S-টা generating set হচ্ছে না, মানে $\text{span}(S)$ হচ্ছে V-এর একটা strict subset.

> Then $\text{span}(S) \subsetneq V$.
> Let $B = \{\vec{v}_1, ..., \vec{v}_n\}$ be a basis.
> Since $\text{span}(B) = V$, $\therefore B \not\subseteq \text{span}(S)$.
> W.l.g., $\vec{v}_1 \notin \text{span}(S)$.
> So $S_1 = S \cup \{\vec{v}_1\}$ is a linearly independent set of size $n + 1$.

এইবার extension theorem লাগাব--

> By extension theorem, we can extend S_1 to a basis C.
> Then $|C| =\geq |S_1| = n + 1 > n = \dim(V)$.
> Thus C is a basis with size more than the dimension of the vector space $(\Rightarrow\Leftarrow)$.
> Hence S must have been a basis.

■

এবার তৃতীয় ফলশ্রুতিটা দেখি। এটা একটু অন্যরকম, কিন্তু এটাও সহজ বুদ্ধিতেই বোঝা যায়। একটা finite dimensional vector space-এর কোনো subspace-এর dimension কখনও vector space-টার dimension-কে ছাড়িয়ে যেতে পারে না। যদি দুটো dimension সমান হয়ে যায়, তবে জানবে যে subspace-টা আসলে পুরো vector space-টাই। একটা উপমা দিয়ে ভাবলেই দেখবে যে, এটা খুবই স্বাভাবিক ব্যাপার--ধরো A হল একটা finite set. আর $B \subseteq A$. তবে বুঝতেই পারছ যে $|B| \leq |A|$ হবেই, আর $|B| = |A|$ হওয়া মানে $B = A$ হতে বাধ্য।

THEOREM

Let V be a finite dimensional vector space. Let U be a subspace of V. Then

- $\dim(U) \le \dim(V)$.
- $\dim(U) = \dim(V) \iff U = V$.

12.1.2 হাতেকলমে

Example 38: Extend the set $S = \{(1,1,1,1),\ (-1,1,-1,1)\}$ to a basis of the real vector space $\mathbb{R}^4$.[5] **(2013.3b)**

SOLUTION: আমরা extension theorem প্রমাণের কায়দাতেই এগোব। সেখানে একটা basis নিয়ে শুরু করেছিলাম। এখানেও তাই করব। সুপরিচিত basis-টাই নেওয়া যাক, $\{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}$. একে একে এই vector-গুলোকে S-এর মধ্যে ঢোকানোর চেষ্টা করব। বোঝাই যাচ্ছে যে, কাজটা দুবারই খালি করা যাবে, কারণ $\dim(\mathbb{R}^4) = 4$ আর $|S| = 2$, যেটা 4-এর চেয়ে 2 কম। প্রথম ধাপে এমন কাউকে ঢোকাতে হবে যে span(S)-এর বাইরে। একটু রাফে কষে দেখলেই দেখবে যে, প্রথম vector-টাই span(S)-এর বাইরে--

Step 1: We notice that $(1,0,0,0) \notin \text{span}(S)$.

⟦ Because:

If $a(1,0,0,0) + b(1,1,1,1) + c(-1,1,-1,1) = (0,0,0,0)$,
then $a+b-c=0,\ b+c=0,$ and $b-c=0$.
So $a=b=c=0$.

⟧

Hence $S_1 = \{(1,0,0,0),(1,1,1,1),(-1,1,-1,1)\}$ is a linearly independent set.

এবার ফের একটু রাফে দেখে নিতে হবে নতুন কোন vector এই S_1-এর span-এর বাইরে পড়ে--

Step 2: We notice that $(0,1,0,0) \notin \text{span}(S_1)$.

⟦ Because:

If

$$a(0,1,0,0)+b(1,0,0,0)+c(1,1,1,1)+d(-1,1,-1,1) = (0,0,0,0),$$

then $b+c-d=0,\ a+c+d=0,\ c-d=0$ and $c+d=0$.
So $a=b=c=d=0$.

⟧

Hence $S_2 = \{(0,1,0,0),(1,0,0,0),(1,1,1,1),(-1,1,-1,1)\}$ is linearly independent.

Also $|S_2| = 4 = \dim(\mathbb{R}^4)$. So S_2 is a basis.

> Since S_2 is a superset of S, we have extended S to a basis.

■

একইরকম আরেকটা অংক, কিন্তু এবার খালি এক ধাপ extend করলেই হবে।

Exercise 39: Extend the set $\{(1,1,-1),\ (2,1,1)\}$ to a basis of $\mathbb{R}^3$.[1] **(2007.12a)** ■

এই অংকগুলো ছোটো ছোটো ছিল, তাই যেভাবে extension theorem-টা প্রমাণ করেছিলাম সেভাবেই করে যেতে পারলাম। প্রতিধাপে আমাদের একটু এদিক ওদিক টুঁ মেরে এমন কোনো vector খুঁজে দেখতে হচ্ছিল, যেটা আমাদের span-এর বাইরে পড়ে। অংকটা একটু বড় হলেই এই টুঁ মেরে দেখাটা কঠিন হয়ে পড়ত। সেরকম ক্ষেত্রে সাধারণতঃ **reduced row echelon form (RREF)** বলে একটা কায়দা ব্যবহার করা হয়। যদিও সেটা সিলেবাসে নেই, কিন্তু কায়দাটা খুবই কাজের, এবং তা দিয়ে অন্য অনেক জিনিসও করা যায়। তাই এই অধ্যায়ের শেষে গিয়ে সেটা একটু আলোচনা করব।
এই যে আমরা ধাপে ধাপে করছিলাম, সেখানে প্রতি ধাপে নতুন করে দেখছিলাম span-এর বাইরে কে আছে। আচ্ছা, একটা শর্টকাট চেষ্টা করা যাক এইভাবে--

> ধরো একটা linearly independent set নিয়ে শুরু করছি, S যার সাইজ 2. কাজ হচ্ছে $\mathbb{R}^4$-এ। তার মানে দুটো vector যোগ করতে হবে S-এর মধ্যে। আমরা এমন দুটো vector বার করলাম $\vec{u},\vec{v}$, যাতে $\vec{u},\vec{v}\notin \text{span}(S)$ এবং $\{\vec{u},\vec{v}\}$ আবার linearly independent-ও বটে। তাহলে $S\cup\{\vec{u},\vec{v}\}$ নিশ্চয়ই $\mathbb{R}^4$-এর একটা basis হবে?

তোমার কী মনে হয়? শর্টকাটটা সর্বদা কাজ করবে? নীচের অংকটা সেটা নিয়েই।

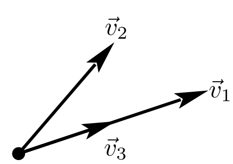
Exercise 40: Let $\{\vec{u}_1,...,\vec{u}_n\}$ be any set of vectors in a vector space V. Let $\{\vec{v}_1,\vec{v}_2\}$ be a linearly independent set such that $\vec{v}_1,\vec{v}_2\notin \text{span}\{\vec{u}_1,...,\vec{u}_n\}$. Then prove or disprove: $\{\vec{v}_1,\vec{v}_2\}\cup\{\vec{u}_1,...,\vec{u}_n\}$ must be a linearly independent set.
■

12.2 Contraction theorem

Extension theorem ছিল ইঁদুরকে বড় করতে করতে শুয়োর করা। এবার শিখব contraction theorem, হাতিকে ছোটো করতে করতে শুয়োর বানানো, মানে generating set-কে ছোটো করতে করতে basis বানানো।

Example 39: Fig 38-এ $\{\vec{v}_1,\vec{v}_2,\vec{v}_3\}$ হল $\mathbb{R}^2$-এর একটা generating set. ঠিক যেন তিন পাহারাদার মিলে $\mathbb{R}^2$-কে পাহারা দিচ্ছে। এদের কাউকে বাদ দিয়েও পুরো $\mathbb{R}^2$-কে পাহারা দেওয়া যায় কি?
SOLUTION: এখানে $\vec{v}_1$ আর $\vec{v}_3$ একই লাইন বরাবর আছে। সুতরাং ওদের একজনকে নিশ্চিন্তে বাদ দেওয়া যায়। ধরো $\vec{v}_1$-কে বাদ দিলাম। আর কিছু বাদ দেওয়া যায় না। লক্ষ করো এখন যে set-টা পড়ে আছে, মানে $\{\vec{v}_2,\vec{v}_3\}$, সেটা একটা

Fig 38



basis. যদি $\vec{v}_1$-এর বদলে $\vec{v}_3$-কে বাদ দিতে তাহলেও একটা basis পেতে, $\{\vec{v}_1, \vec{v}_2\}$. ■

যদি তোমাকে একটা generating set দিই যেটা finite, তবে এই কায়দাটা সব সময়েই খাটবে। তুমি অনাবশ্যকগুলোকে একটা একটা করে বেছে বেছে ফেলে দেবে। শেষমেশ একটা basis পড়ে থাকতে বাধ্য।
অর্থাৎ কিনা হাতি ছোটো হতে শুয়োর হয়!

> **THEOREM**
>
> If V is any vector space, and $G \subseteq V$ is any finite generating set, then there must be a subset $B \subseteq G$ such that B is a basis of V.

Example 40: Prove that there exists a basis for each finite dimensional vector space.[5] **(2022, 2009)**

SOLUTION: এই অংকের ভাষাটায় একটু সমস্যা আছে। "Dimension" মানে হল "basis-এর সাইজ"। অতএব "finite dimensional vector space" বলতে বোঝায় এমন vector space, যার একটা finite basis আছে। সেটাই যদি শর্ত হিসেবে ধরে নেওয়া হয়, তবে basis-এর existence আর নতুন করে দেখানোর কী আছে?
আসলে অংকটায় বলতে চেয়েছে "finitely generated vector space", অর্থাৎ এমন একটা vector space, যার অন্ততঃ একটা generating set আছে, যেটা finite.

> By definition, the trivial vector space has ϕ as its basis.

সুতরাং খালি nontrivial-দের জন্য প্রমাণটা করতে পারলেই হবে।

> Let V be a nontrivial vector space over a field F.
> Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be a generating set.
> If it is linearly independent, then it is a basis, and the proof is done.

নইলে contraction theorem লাগাব--

> Otherwise, $\exists a_1, ..., a_n \in F$, not all zero, such that
>
> $$\sum_1^n a_i \vec{v}_i = \vec{0}.$$
>
> W.l.g., let $a_n \neq 0$. Then
>
> $$\vec{v}_n = \sum_1^{n-1} \left(-\tfrac{a_i}{a_1}\right) \vec{v}_i \in \text{span}\{\vec{v}_1, ..., \vec{v}_{n-1}\}.$$
>
> Let $S_1 = \{\vec{v}_1, ..., \vec{v}_{n-1}\}$.
> Then $\text{span}(S_1) = \text{span}(S) = V$, and $|S_1| = n - 1$.
> If S_1 is linearly independent, then we are done.

> Otherwise, we can apply the same method to S_1 to get smaller and smaller generating sets. This process can continue for at most $n-1$ steps because S is finite.
> So we always get a subset of S such that it is a basis of V.

■

12.2.1 ফলশ্রুতি

আমরা extension theorem দিয়ে যেমন পেয়েছিলাম--

basis হল maximal linearly independent set,

সেই একই পথে contraction theorem থেকে পাব--

basis মানে **minimal generating set**.

অর্থাৎ যদি এমন একটা generating set পাও, যার থেকে একটাও vector বাদ দিলে সেটা আর generating set থাকছে না, তবে সেই generating set-টা অতি অবশ্যই একটা basis হতে বাধ্য। বিপরীতপক্ষে basis হল এমন একটা generating set, যার থেকে একটাও vector বাদ দিলে সেটা আর generating থাকবে না।

THEOREM

A subset of a finite dimensional vector space is a basis if and only if it is a minimal generating set.

এই ব্যাপারটাকেই dimension ব্যবহার করে লিখলে হয়--

THEOREM

Let V be a vector space of dimension n. Let $S \subseteq V$ be a generating set. Then

- $|S| \geq n$.
- $|S| = n \iff S$ is a basis of V.

Exercise 41: Show that $\mathbb{R}^3$ cannot be spanned with two vectors. **(2021)** ■

12.2.2 হাতেকলমে

Contraction theorem যেভাবে প্রমাণ করলাম সেটা হাতে-কলমে করা কিন্তু সহজ নাও হতে পারে। কারণ প্রতি ধাপে দেখতে হবে যারা পড়ে আছে তারা linearly independent কিনা, এবং উত্তরটা যদি "না" হয়, তবে খুঁজে বার করতে হবে

কে পরের পয়সায় বসে খাচ্ছে। যদি তালিকাটা ছোটো হয়, তবে এই দুটো কাজ চোখে দেখেই করে ফেলা যায়, কিন্তু বড় অংকের ক্ষেত্রে কায়দা করে না এগোলে মুস্কিল। এখানেও সেই reduced row echelon form (RREF) ব্যবহার করে সহজে এগোনো যায়। সেটা আমরা পরে বলব। আপাততঃ নীচের অংকটা যথেষ্ট ছোটো, তাই চোখে দেখেই করে ফেলব।

Example 41: Find a linearly independent subset T of the set

$$S = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\},$$

where

$$\alpha_1 = (1, 2, -1), \quad \alpha_2 = (-3, -6, 3), \quad \alpha_3 = (2, 1, 3), \quad \alpha_4 = (8, 7, 7) \in \mathbb{R}^3$$

which spans the same space as S.[5] **(2006.12c)**

SOLUTION: এখানে তালিকাটা একেবারেই ছোট্টো। তাই চোখে দেখেই বোঝা যায় কে কার পয়সায় বসে খাচ্ছে।

> We observe that $\alpha_2 = -3\alpha_1$.

এইটা চোখে দেখেই বুঝে গেলাম। তার মানে α_1, α_2-এর মধ্যে যেকোনো একজনকে বিদায় দেওয়া যায়। ধরো α_1-কে বিদায় দিলাম--

> So $\mathrm{span}\{\alpha_1, \alpha_3, \alpha_4\} = \mathrm{span}(S)$.

এবার আরো একটু চোখ পাকিয়ে খুঁজতে হবে--

> Again, $\alpha_4 = 2\alpha_1 + 3\alpha_3$.

এটা খুঁজে পাওয়া কিন্তু তত সহজ ছিল না। যাই হোক পেয়ে যখন গেছি--

> So $\mathrm{span}\{\alpha_1, \alpha_3\} = \mathrm{span}(S)$.

এবার মোটে দুটো vector পড়ে আছে, সুতরাং ওরা পরস্পরের multiple কি না দেখলেই চলবে--

> Now α_1 and α_3 are not multiples of each other. So $\{\alpha_1, \alpha_3\}$ is linearly independent.
>
> So $T = \{\alpha_1, \alpha_3\}$ is a basis of $\mathrm{span}(S)$.
>
> Thus $T \subseteq S$ satisfies the required conditions.

■

DAY 13
Algebra of subspaces (part 1)

যদি তোমাকে কিছু সংখ্যা দিই, তবে তুমি যেমন তাদের যোগ বিয়োগ ইত্যাদি করে আরেকটা সংখ্যা বানাতে পারো, তেমনি বিভিন্ন subspace-কে মিলিয়ে জুলিয়ে নতুন নতুন subspace বানানো যায়। এরকম কয়েকটা কায়দা আলোচনা করব এবার।

13.1 Intersection

প্রথম কায়দা হল subspace-দের intersection নেওয়া। তবে যে জিনিসটা পাবে সেটাও ফের একটা subspace হতে বাধ্য। চট্ করে একটা উদাহরণ দেখে নিই।

Example 42: $\mathbb{R}^3$-এর মধ্যে দুটো subspace হল x-axis আর y-axis. এদের intersection-টা কি একটা subspace? এর dimension কত?

SOLUTION:এদের intersection হল $\{(0,0,0)\}$, যেটা অবশ্যই একটা subspace. এর dimension হল 0. ■

এই উদাহরণটা ছিল একবারে একটা চরম উদাহরণ, কারণ যেকোনো দুটো subspace-এর মধ্যেই $\vec{0}$ তো থাকবেই, তাই $\vec{0}$-টা ওদের intersection-এও থাকবে। এই উদাহরণে সেটাই ছিল intersection-এর একমাত্র সদস্য।

Exercise 42: আবার $\mathbb{R}^3$-এর মধ্যেই কাজ করব। এবার subspace দুটো নেব

$$U = \{(x,y,z) \;:\; x+y+z=0\} \text{ এবং } W = \{(x,y,z) \;:\; y=z\}.$$

এদের intersection-এর মধ্যে একটা nonzero vector দাও। Intersection-টাকে দেখতে কী রকম হবে--একটা বিন্দু, নাকি একটা লাইন, নাকি একটা plane, বা পুরো $\mathbb{R}^3$? ■

Subspace-দের intersection যে ফের একটা subspace-ই হবে, সেই কথাটা গুছিয়ে লিখে প্রমাণ করা যাক।

> THEOREM
>
> Arbitrary intersection of subspaces of a vector space is again a subspace.

এখানে আমরা arbitrary intersection বলেছি। এর মানে কতগুলো subspace-এর intersection নিচ্ছি, তাতে কিছু এসে যায় না, দুটো, তিনটে চারটে, finite, infinite ইত্যাদি যতগুলো খুশি subspace-এরই intersection নাও না কেন, সর্বদা ফের একটা subspace-ই পাবে।
নীচের অংকে finite-সংখ্যক subspace-এর জন্য প্রমাণটা চেয়েছে। একই যুক্তি অবশ্য infinite-সংখ্যকের বেলাতেও খাটবে।

Example 43: Prove that the intersection of any finite number of subspaces of a vector space V over a field F is a subspace of V.[2] **(2009.10a)**

SOLUTION:

> Let $W_1, ..., W_n$ be subspaces of V.
> Shall show that $W = W_1 \cap \cdots \cap W_n$ is a subspace of V.

আমরা আগেই শিখেছি একটা subset-কে কী করে subspace দেখাতে হয়। প্রথমে দেখাতে হয় যে, ওর মধ্যে $\vec{0}$ আছে, তারপর দেখাতে হয় যে, $\vec{w}, \vec{z}$ যদি ওর মধ্যে থাকে, তবে যেকোনো $a, b \in F$-এর জন্যই $a\vec{w} + b\vec{z}$-ও ওর মধ্যে থাকবে। প্রথম শর্তটা পরীক্ষা করা খুবই সহজ--

> $\because \forall i \quad \vec{0} \in W_i$
> $\therefore \vec{0} \in W.$

এবার দ্বিতীয় শর্তটা দেখি--

Shall show

$\forall \vec{w}, \vec{z} \in W \;\; \forall a, b \in F \quad a\vec{w} + b\vec{z} \in V.$

$\forall \vec{w}, \vec{z}$ Take any $\vec{w}, \vec{z} \in W.$

$\forall a, b$ Take any $a, b \in F.$

Then $\forall i \quad \vec{w}, \vec{z} \in W_i.$

$\therefore \forall i \quad a\vec{w} + b\vec{z} \in W_i.$

$\therefore a\vec{w} + b\vec{z} \in W,$ as required.

Hence W is a subspace of V.

■

13.2 Union

যেহেতু subspace-দের intersection ফের একটা subspace হল, তাই উৎসাহের চোটে তুমি মনে করে বসতে পারো যে, subspace-দের union-ও বুঝি একটা subspace হবে। তা যে নাও হতে পারে, সেটা নীচের অংকটা দেখলেই বুঝবে।

Example 44: $\mathbb{R}^2$-এর মধ্যে দুটো subspace হল x-axis আর y-axis. এদের union কি একটা subspace?

SOLUTION: না, কারণ $(1,0)$ এবং $(0,1)$ দুজনেই এই union-এর মধ্যে আছে, কিন্তু $(1,0)+(0,1)=(1,1)$ মোটেই নেই। ■

Exercise 43: $\mathbb{R}^3$-এর মধ্যে দুটো subspace হল x-axis এবং xy-plane। এদের union কি একটা subspace? ■

এই দুটো অংক থেকে যা শিখলে সেটার ভিত্তিতেই নীচের অংকটা তৈরী।

Exercise 44: Let U, W be both subspaces of a vector space V. Then show that $U \cup W$ is a subspace if and only if either $U \subseteq W$ or $W \subseteq U$.
HINT: এখানে if অংশটা দেখানো খুবই সহজ, $U \subseteq W$ বা $W \subseteq U$ হলে যে $U \cup W$ একটা subspace হবে, সেই দিকটা। Only if অংশটার জন্য ধরে নাও $U \cup W$ একটা subspace এবং $U \not\subseteq W$. এ থেকে দেখাও $W \subseteq U$ হতে বাধ্য। ■

13.3 Sum

দুটো subspace-এর union যে ফের একটা subspace নাও হতে পারে, তার কারণ এই যে, যদি তুমি দুটো subspace থেকে একটা করে vector নাও, তবে যাদের যোগফলটা union-এর বাইরে চলে যেতে পারে। সেই সমস্যা এড়ানোর একটা পথ হল দুটো subspace-এর vector-দেরকে যতভাবে যোগ করা যায়, সেই সব যোগফলের set-টা নিয়ে কাজ করা। মানে S আর T যদি দুটো subspace হয়, তবে আমরা S-এর থেকে যেকোনো vector নিচ্ছি, $\vec{s}$, আবার T-এর থেকে যেকোনো vector নিচ্ছি $\vec{t}$. এদের যোগ করে পাচ্ছি $\vec{s}+\vec{t}$. এরকম সব যোগফলের set-টাকে বলব S এবং T-এর **sum,** এবং লিখব--

$$S + T = \{\vec{s} + \vec{t} \;:\; \vec{s} \in S,\; \vec{t} \in T\}.$$

এই $S+T$ ফের একটা subspace-ই হবে। একটা উদাহরণ দেখা যাক।

Example 45: ধরো $\mathbb{R}^3$-এর মধ্যে x-axis আর y-axis নিলাম। এদের নাম দাও, যথাক্রমে S আর T. তবে $S+T$ কী হবে?

SOLUTION: এখানে S-এর সদস্যদের চেহারা হল $(x,0,0)$-র মত, আর T-এর সদস্যরা দেখতে $(0,y,0)$-র মত। তাই $S+T$-এর সদস্যদের দেখতে হবে $(x,y,0)$-এর মত। এবং এরকম যেকোনো vector-ই $S+T$-এর মধ্যে থাকবে। তাই $S+T$ এখানে xy-plane-টা হবে। ■

সংজ্ঞাটা গুছিয়ে লিখে রাখি।

> DEFINITION: **Sum of vector spaces**
>
> If S,T are two subspaces of a vector space, then we define
>
> $$S+T=\{\vec{s}+\vec{t}\ :\ \vec{s}\in S,\ \vec{t}\in T\}.$$

Exercise 45: If $S=\{(x,y,z)\ :\ x+y+z=0\}$ and $T=\{(x,y,z)\ :\ x+y-z=0\}$ then find $S+T$. ■

Example 46: Let U and W be two subspaces of a vector space V over field F and

$$U+W=\{u+w\ :\ u\in U,\ w\in W\}.$$

Prove that $U+W$ is the smallest subspace of V containing the subspaces U and W.[5] **(2013.3a)**

SOLUTION: আমরা তিন ধাপে এগোব--প্রথম ধাপে দেখাব যে $U+W$ সত্যিই V-এর একটা subspace. দ্বিতীয় ধাপে দেখাব যে $U,W\subseteq U+W$. আর তৃতীয় ধাপে দেখাব যে, $U+W$-এর চেয়ে ছোটো কোনো subspace-এর মধ্যে U এবং W দুজনেই একইসঙ্গে থাকতে পারে না।

<u>Step 1:</u> To show that $U+W$ is a subspace of V.

Clearly $\vec{0}\in U+W(\because \vec{0}\in U,\ \vec{0}\in W,\ \vec{0}+\vec{0}=\vec{0})$.

Shall show

$\forall a_1,a_2\in F\quad \forall \vec{v}_1,\vec{v}_2\in U+W\quad a_1\vec{v}_1+a_2\vec{v}_2\in U+W.$

$\forall a_1,a_2$ Take any $a_1,a_2\in F$.

$\forall \vec{v}_1,\vec{v}_2$ Take any $\vec{v}_1,\vec{v}_2\in U+W$.

Then $\vec{v}_1=\vec{u}_1+\vec{w}_1$ for some $\vec{u}_1\in U$ and $\vec{w}_1\in W$.

Similarly, $\vec{v}_2=\vec{u}_2+\vec{w}_2$ for some $\vec{u}_2\in U$ and $\vec{w}_2\in W$.

So

$$a_1\vec{v}_1+a_2\vec{v}_2 = (a_1\vec{u}_1+a_2\vec{u}_2)+(a_1\vec{w}_1+a_2\vec{w}_2),$$

using distributivity and commutativity. Now, since U, W are subspaces of V, hence

$$a_1\vec{u}_1 + a_2\vec{u}_2 \in U \text{ and } a_1\vec{w}_1 + a_2\vec{w}_2 \in W.$$

So $a_1\vec{v}_1 + a_2\vec{v}_2 \in U + W$, as required.

এবার দ্বিতীয় ধাপ, যার মধ্যে কোনো জটিলতা নেই।

Step 2: Shall show that $U \subseteq U + W$ and $W \subseteq U + W$.

Take any $\vec{u} \in U$. Since W is a subspace, hence $\vec{0} \in W$.

Hence $\vec{u} = \vec{u} + \vec{0} \in U + W$.

So $U \subseteq U + W$.

Similarly, $W \subseteq U + W$.

তৃতীয় ধাপে কী করব বুঝে নাও। আমরা দেখাব যে $U + W$-এর চেয়ে ছোটো এমন কোনো subspace পাবে না, যার মধ্যে U, W দুজনেই আছে। মানে ঘুরিয়ে বললে, যদি T হয় এমন কোনো subspace যাতে $U, W \subseteq T$, তবে T নিশ্চয়ই $U + W$-এর থেকে বড়, অর্থাৎ $U + W \subseteq T$. এই কথাটাকে লিখে নিয়ে শুরু করি--

Step 3: Shall show that if T is a subspace of V such that $U, W \subseteq T$, then we must have $U + W \subseteq T$.

Let T be such a subspace.

Take any $\vec{v} \in U + W$.

Then $\vec{v} = \vec{u} + \vec{w}$ for some $\vec{u} \in U$ and $\vec{w} \in W$.

Thus, $\vec{u}, \vec{w} \in T$.

Since T is a subspace, so $\vec{u} + \vec{w} \in T$.

So $\vec{v} \in T$.

So $U + W \subseteq T$, as required.

■

আমরা দেখেছিলাম যে, দুটো subspace দেওয়া থাকলে তাদের intersection ফের একটা subspace হয়। কিন্তু union-এর বেলায় সে গ্যারান্টিটা ছিল না। অবশ্য তার বদলে আমরা $S + T$ নিয়ে কাজ করতে পারলাম। স্বভাবতঃই প্রশ্ন ওঠে S^c-ও একটা subspace কিনা। এখানে উত্তরটা যে "না" হবে, সেটা বোঝাই যায়, কারণ যেকোনো subspace-র মধ্যেই $\vec{0}$ থাকতে হবে। যেহেতু $\vec{0} \in S$, তাই S^c-এর মধ্যে আর $\vec{0}$ থাকে কী করে? যদি S একটা subspace হয়, তবে "S^c-জাতীয় একটা subspace" তৈরীর চেষ্টা বিভিন্নভাবে করা যায়। এরকম দুটো পথ আমরা এখানে আলোচনা করব, direct complement এবং quotient.

13.4 Direct complement

এর জন্য প্রথমে S-এর যেকোনো একটা basis নাও। তারপর সেটাকে পুরো V-এর একটা basis-এ extend করো। ধরো S-এর basis-টা ছিল B, আর extend করার জন্য বাড়তি যা লাগাতে হল সেটা হল C. তবে span(C)-কে বলব S-এর একটা **direct complement**. এখানে "একটা" বলার কারণ এই যে, আমরা S-এর basis-টা নানাভাবে নিতে পারতাম, এবং extend করাটাও নানাভাবে করতে পারতাম। সবক্ষেত্রেই যে span(C) একই আসবে এমন কোনো কথা নেই।

Example 47: ধরো আমাদের vector space-টা হল $\mathbb{R}^2$. তার মধ্যে S হল x-axis-টা। এর কয়েকটা direct subspace বার করো।

SOLUTION: S-এর একটা basis হল $\{(1,0)\}$. একে আমরা extend করে পুরো $\mathbb{R}^2$-র এই basis-টা পেতে পারি, $\{(1,0),(0,1)\}$. এখানে বাড়তি এল $(0,1)$. এর span-টা হল y-axis. সুতরাং y-axis হল x-axis-এর একটা direct complement.

চাইলে আমরা $\mathbb{R}^2$-র basis-টা এরকমও নিতে পারতাম, $\{(1,0),(1,1)\}$. সেক্ষেত্রে direct complement-টা হত span$\{(1,1)\}$. আশা করি বুঝতে পারছ যে, origin দিয়ে যাওয়া যেকোনো লাইনই (x-axis নিজে ছাড়া) হবে x-axis-এর একটা direct complement. ■

Exercise 46: নীচের প্রতিক্ষেত্রে S হল $\mathbb{R}^3$-এর একটা subspace. তোমার কাজ হল S-এর দুটো direct complement বার করা। যদি কোনোক্ষেত্রে দুটো direct complement না থাকে, সেটা বলতে হবে।

1. $S =$ x-axis.
2. $S =$ xy-plane.
3. $S = \text{span}\{(1,1,1)\}$.
4. $S = \{\vec{0}\}$.
5. $S = \mathbb{R}^3$.

■

13.4.1 অংকের ভাষায়

Direct complement-এর সংজ্ঞাটা সাধারণতঃ তার একটা ধর্মকে ব্যবহার করে দেওয়া হয়। লক্ষ করো, যদি V-এর মধ্যে S একটা subspace হয়, আর T হয় তার কোনো একটা direct complement, তবে দুটো জিনিস হতে বাধ্য--

- এক, $S \cap T = \{\vec{0}\}$,
- আর দুই, $S + T = V$.

বিপরীতপক্ষে, এই দুটো শর্ত পালন করে যা-ই T তুমি নাও না কেন, সেটা S-এর একটা direct complement হবেই। তাই এই দুটো শর্তকেই direct complement-এর সংজ্ঞা হিসেবে ব্যবহার করা যায়--

DEFINITION: **Direct complement**

Let S be a subspace of a vector space V. Then by **direct complement** of S in V we understand any subspace T of V satisfying these conditions

1. $S \cap T = \{\vec{0}\}$,
2. $S + T = V$.

We write this as $S \oplus T = V$.

আমরা $S \oplus T$-কে পড়ি S আর T-এর **direct sum**.
এই সংজ্ঞা থেকে একটা জিনিস স্পষ্ট--যদি T হয় S-এর একটা direct complement, তবে S-ও আবার হবে T-এর একটা direct complement. তবে এই সংজ্ঞা থেকে এটা বোঝা যাচ্ছে না যে, যেকোনো subspace-এরই অন্ততঃ একটা direct complement আছে। সেটা আবার basis দিয়ে ভাবলে সহজে বোঝা যায়।

Exercise 47: সরাসরি সংজ্ঞাটা ব্যবহার করে বলো নীচের বক্তব্য দুটো $\mathbb{R}^3$-এর ক্ষেত্রে ঠিক নাকি ভুল--

1. x-axis এবং y-axis পরস্পরের direct complement.
2. xy-plane এবং xz-plane পরস্পরের direct complement.

■

Exercise 48: If S and T are direct complements in V, then show that for every $\vec{v} \in V$ there are *unique* $\vec{s} \in S$, $\vec{t} \in T$ such that $\vec{v} = \vec{s} + \vec{t}$. ■

Exercise 49: Let S and T be subspaces of V such that for every $\vec{v} \in V$ there are *unique* $\vec{s} \in S$, $\vec{t} \in T$ with $\vec{v} = \vec{s} + \vec{t}$. Then show that S and T must be direct complements. ■

13.5 Quotient

এবার যেটা বলতে চলেছি, সেটাও কতকটা একটা subspace-এর complement-এর মতন ব্যাপার, তবে সামান্য একটু বেশী জটিল। প্রথমে ছবি বোঝা যাক।

Example 48: $\mathbb{R}^3$-এর মধ্যে xy-plane-টাকে S হিসেবে নেওয়া যাক। যদি $\mathbb{R}^3$-কে একটা ঘর বলে ভাবি, তবে S-টা

যেন তার মেঝে (Fig 39)। অবশ্য এ ঘরে দেওয়াল বা ছাদের বালাই নেই, এবং ঘরটা মেঝের নীচেও বিস্তৃত। এবার যদি ঘরটার মধ্যে একটা বিশাল মঞ্চ বানানো যায়, তবে সেই মঞ্চের উপরটাও যেন আরেকটা মেঝের মতই হবে (Fig 40), খালি আসল মেঝের খানিকটা উপরে অবস্থিত। পুরো $\mathbb{R}^3$-কেই এরকম অসংখ্য মঞ্চ পরপর সাজিয়ে তৈরী বলে ভাবা যায় (Fig 41)। চাইলে তুমি এগুলোকে যেন বইয়ের একেকটা পাতা বলে ভাবতে পারো, যারা পরস্পরের সঙ্গে সমান্তরালভাবে সাজানো। এর মধ্যে একটাই খালি আমাদের S-টা। এবার যে পাতাটা 1 উচ্চতায় আছে, সেখান থেকে যা খুশি একটা vector নিলাম, $\vec{u}$. আবার যে পাতাটা 2 উচ্চতায় আছে, সেখান থেকেও একটা vector নিলাম, $\vec{v}$. তবে $\vec{u} + \vec{v}$-টা কোন পাতায় থাকতে বাধ্য?
SOLUTION: যেহেতু $\vec{u}$ আছে 1 উচ্চতায়, তাই $\vec{u}$ হবে $(a, b, 1)$ চেহারার। একইভাবে 2 উচ্চতার পাতার $\vec{v}$-কে দেখতে $(c, d, 2)$-এর মত। $\vec{u} + \vec{v} = (a + c, b + d, 3)$, যেটা থাকবে 3 উচ্চতার পাতায়। ■

লক্ষ্য করো, $\vec{u}$ আর $\vec{v}$ নিজের নিজের পাতার কোথায় আছে সেটা বড় কথা নয়। তুমি $\vec{u}$-কে 1 উচ্চতার পাতায় এবং $\vec{v}$-কে 2 উচ্চতার পাতার যেখানেই নাও না কেন, $\vec{u} + \vec{v}$ ঠিক 3 উচ্চতার পাতাতেই থাকতে বাধ্য। তাই আমরা বলতে পারি যেন

Fig 39

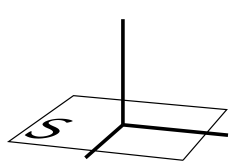


Fig 40

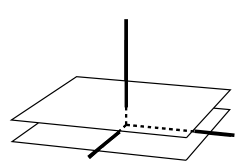
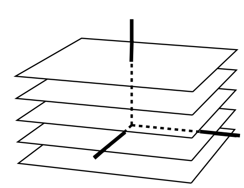


Fig 41



1 উচ্চতার পাতা + 2 উচ্চতার পাতা = 3 উচ্চতার পাতা।

এর মধ্যে দুটো দিকই ধরা আছে--

- এক, যদি বাঁদিকের পাতা দুটো থেকে একটা করে vector নাও, তবে তাদের যোগফল ডানদিকের পাতায় থাকবে,
- এবং দুই, ডানদিকের পাতার যেকোনো vector-কেই তুমি দুভাগে ভেঙে লিখতে পারবে, যাতে একটা ভাগ থাকে 1 উচ্চতার পাতায়, আর অন্যটা থাকে 2 উচ্চতার পাতায়, (যেমন $(a, b, 3)$-কে ভেঙে লেখা যায় $(a, b, 1) + (0, 0, 2)$ আকারে[3])

দুটো পাতাকে যেমন যোগ করা যায়, তেমনি চাইলে একটা পাতাকে কোনো সংখ্যা দিয়ে গুণও করা যায়। যেমন নীচের উদাহরণে।

Example 49: ধরো আমরা 1 উচ্চতার পাতাটাকে 2 দিয়ে গুণ করলাম। তবে কী পাবে? অর্থাৎ ওই পাতার সব vector-কেই যদি দ্বিগুণ করো, তবে সেই সব দ্বিগুণ vector-দের set-টা কী হবে?

SOLUTION: 1 উচ্চতার পাতার প্রতিটা vector-ই দেখতে $(x, y, 1)$-এর মত, তাদের দ্বিগুণ করলে পাবে $(2x, 2y, 2)$-এর মত দেখতে vector. এরা সকলেই 2 উচ্চতার পাতায় আছে। বিপরীতপক্ষে 2 উচ্চতার পাতার সব vector-কেই তুমি 1 উচ্চতার পাতার কোনো vector-এর দ্বিগুণ বলে ভাবতে পারো (যেমন $(4, 3, 2)$-কে ভাবতে পারো $2 \times \left(2, \frac{3}{2}, 1\right)$ বলে)। ■

এই একই যুক্তিতে এটাও দেখা যায় যে, যেকোনো $a, b, r \in \mathbb{R}$-এর জন্যই--

a উচ্চতার পাতা + b উচ্চতার পাতা = $(a + b)$ উচ্চতার পাতা,

এবং

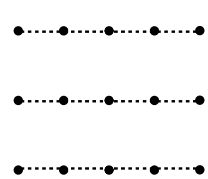
$r \times$ (a উচ্চতার পাতা) = (ra) উচ্চতার পাতা।

অর্থাৎ পাতাগুলো নিজেরাই যেন এককেটা vector-এর মত আচরণ করছে। এই পাতাগুলো দিয়ে যে নতুন vector space-টা তৈরী হচ্ছে, সেটাকে লিখি V/S আর মুখে বলি **quotient space** of $V$by S. এখানে "quotient" শব্দটার মধ্যে "ভাগ" -এর ধারণা আছে, আর V/S চিহ্নটার মধ্যেও একটা ভাগচিহ্ন আছে। এর পিছনের কারণটা একটা সহজ উদাহরণ দিয়ে বলি।

Example 50: Fig 42-এ পনেরোটা বিন্দু আছে, তাদেরকে লাইনে লাইনে সাজানো আছে, প্রতিটা লাইনে পাঁচটা করে বিন্দু আছে। এরকম কটা লাইন আছে? উত্তর হল $3 = 15/5$. এই লাইনগুলো যেন এককেটা পাতা। প্রতিটা লাইনেই যেহেতু সমান সংখ্যক বিন্দু ছিল, তাই লাইন গোণার সময়ে ভাগ করা গেল। Vector space-এর বেলাতেও পাতাগুলো সব একইরকম দেখতে ছিল, তাই পাতাগুলোর set-টা যেন পুরো vector space-কে পাতা দিয়ে ভাগ করা হচ্ছে। ■

[3] আরো অসংখ্যভাবেও কাজটা করা যায়, যেমন ধরো $(a, 0, 1) + (0, b, 2)$.

Fig 42



অবশ্যই এটা একটা উপমা মাত্র, notation-টার সঙ্গে বন্ধুত্ব করে নেবার জন্য।
এই যে পাতার গল্পটা বললাম, এবার সেটাকে অংকের ভাষায় প্রকাশ করব।

13.5.1 অংকের ভাষায়

প্রথমে একটা নতুন notation শিখে নিই--একটা vector space-এর কোনো subset-এর সঙ্গে একটা vector যোগ করা। ধরো V হল আমাদের vector space. এর মধ্যে একটা subset হল S আর একটা vector হল $\vec{v}$. এবার যদি S-এর প্রতিটা vector-এর সঙ্গে $\vec{v}$ যোগ করো, তবে যে নতুন set-টা পাবে, তাকে লিখব $\vec{v}+S$, মানে--

$$\vec{v}+S=\{\vec{v}+\vec{s}\ :\ \vec{s}\in S\}.$$

ছবি দিয়ে ভাবলে S-কে যেন $\vec{v}$ দিকে বরাবর সরিয়ে $\vec{v}+S$ পাওয়া যায়।

Example 51: যদি $\mathbb{R}^3$-এর মধ্যে S হয় xy-plane-টা আর $\vec{v}$ হয় $(a,b,1)$ চেহারার কোনো vector, তবে $\vec{v}+S$ হবে 1 উচ্চতার পাতাটা। এখানে একবার $\vec{v}_1(0,0,1)$ এবং একবার $\vec{v}_2(1,2,1)$ নিয়ে দেখাও যে $\vec{v}_1+S=\vec{v}_2+S$ হচ্ছে।
SOLUTION: $\vec{v}_1+S$-এর মধ্যে যাই vector নেবে, সেটা $(0,0,1)+(x,y,0)=(x,y,1)$ চেহারার হবে। এখন $(x,y,1)=(1,2,1)+(x-1,y-2,0)$. যেহেতু $(x-1,y-2,0)\in S$, তাই vector-টা $\vec{v}_2+S$-এর মধ্যেও থাকবে। সুতরাং $\vec{v}_1+S\subseteq\vec{v}_2+S$ হচ্ছে। একইভাবে অন্য দিকটাও দেখানো যাবে। ■

Exercise 50: এই তিনজনের মধ্যে দুজন আসলে একই জিনিস, কারা?

$$S,\quad (1,0,1)+S,\quad (1,1,0)+S.$$

■

যদি S-টা V-এর মধ্যে হেলানো কিছু হত, সেক্ষেত্রেও আমরা একইরকম জিনিস করতে পারতাম। সেক্ষেত্রে পাতাগুলো সবাই একইভাবে হেলানো হত (Fig 43)। সব সময়ে যে S-কে একটা পাতার মতই হতে হবে (মানে 2-dimensional-ই হতে হবে) এমন কোনো কথা নেই। যদি 1-dimensional হত, তবে S হত একটা লাইনের মত, এবং প্রতিটা $\vec{v}+S$ হত তার সমান্তরাল একটা করে লাইন। সব মিলিয়ে দেখাবে যেন এক আঁটি খড় Fig 44। কিন্তু দেখতে যেমনই হোক, সর্বদাই নীচের theorem-টা খাটতে বাধ্য--

Fig 43

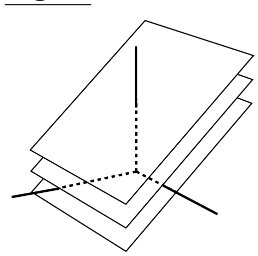
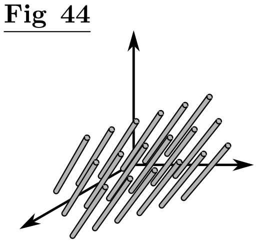


Fig 44



THEOREM

Let V be a vector space over a field F. Let S be any subspace in it. Then for any $\vec{u}, \vec{v} \in V$ and any $r \in F$ we have

1. $(\vec{u}+S)+(\vec{v}+S) = (\vec{u}+\vec{v})+S$,
2. $r(\vec{v}+S) = (r\vec{v})+S$.

সুতরাং $\vec{v}+S$-রা সবাই মিলে একটা vector space তৈরী করে। সেটাকে যে V/S বলে সে তো বলেইছি। সংজ্ঞার আকারে লিখে রাখি।

DEFINITION: Quotient space

Let V be a vector space. Let S be a subspace in V. Then the **quotient space** V/S is the vector space with the set

$$\{\vec{v}+S \ : \ \vec{v} \in V\}$$

along with vector addition defined as

$$(\vec{u}+S)+(\vec{v}+S) = (\vec{u}+\vec{v})+S, \quad \vec{u}, \vec{v} \in V,$$

and scalar multiplication defined as

$$r(\vec{v}+S) = (r\vec{v})+S, \quad r \in F, \vec{v} \in V.$$

DAY 14
Algebra of subspaces (part 2)

14.1 Dimension বার করা

গতকাল আমরা পুরোনো subspace-দের মিলিয়ে নানারকম নতুন subspace বানানো শিখেছি। আজকে দেখব এদের dimension সম্বন্ধে কী বলা যায়। শুরুতেই একটা ব্যাপার বলে নিই, যেটা extension theorem থেকে আসে। আমরা আগেই দেখেছি যে, যদি V একটা vector space হয়, যার একটা set হল S, তবে $\dim(S) \leq \dim(V)$ হতে বাধ্য, কারণ S-এর যেকোনো basis-কে extend করে V-এর basis বানানো যায়। শুধু তাই নয়, $\dim(S) = \dim(V)$ হবে একমাত্র তখনই যখন $S = V$ হবে। একইভাবে যদি V-এর দুটো subspace নিই, S আর T, যেখানে $S \subseteq T$, তবে $\dim(S) \leq \dim(T)$ হতে বাধ্য, এবং $\dim(S) = \dim(T)$ হওয়া মানে $S = T$ হওয়া। যুক্তিটা এখানেও একই--আমরা শুরু করব S-এর একটা basis নিয়ে, এবং extend করার সময়ে খালি T-এর vector-দের দিয়েই extend করব। তাহলে extend করে T-এর একটা basis পাব।
এই ব্যাপারটা মাথায় রেখে আজকের আলোচনাটা শুরু করা যাক।

14.1.1 Intersection

যদি S আর T দুজনেই subspace হয়, তবে $\dim(S \cap T)$ সম্বন্ধে খালি এটুকুই বলা যায় যে, ওটা $\dim(S)$ আর $\dim(T)$, দুজনের চেয়েই কম বা সমান হবে। সেটা বোঝাই যাচ্ছে, কারণ $S \cap T \subseteq S$ এবং $S \cap T \subseteq T$.

> THEOREM
>
> If S, T are two finite dimensional subspaces of V, then
>
> $$0 \leq \dim(S \cap T) \leq \dim(S), \dim(T).$$

Example 52: চট্ করে এমন একটা উদাহরণ দাও, যেখানে $\dim(S) = 3$, $\dim(T) = 5$ আর $\dim(S \cap T) = 2$.

SOLUTION: এরকম উদাহরণ বানানোর জন্য শুরু করো $\mathbb{R}^n$ নিয়ে (n কত সেটা এক্ষুণি ঠিক করব)। এর সুপরিচিত basis-টা তো জানোই, $\{\vec{e}_1, ..., \vec{e}_n\}$. এখানে আমরা সেটা নিয়েই কাজ করব।
S, T আর $S \cap T$-এর মধ্যে সবচেয়ে ছোটো হল $S \cap T$. প্রথমে সেটা বানাও। যেহেতু ওর dimension চাই 2, তাই ওকে নাও $\operatorname{span}\{\vec{e}_1, \vec{e}_2\}$. এবার লক্ষ করো, S-এর dimension এর চেয়ে ঠিক 1 বেশী, তাই $S = \operatorname{span}\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ নাও। ওদিকে T-এর dimension হল 3 বেশী। তাই $S = \operatorname{span}\{\vec{e}_1, \vec{e}_2, \vec{e}_4, \vec{e}_5, \vec{e}_6\}$ নিলেই হবে। এখানে আমাদের সবচেয়ে বেশী $\vec{e}_6$ অব্ধি যেতে হল। তাই $n = 6$ নিলেই হবে (তার চেয়ে বেশী নিলেও অসুবিধা ছিল না)। ■

Exercise 51: Give an example of subspaces S and T such that $\dim(S) = \dim(T) = 6$ and $\dim(S \cap T) = 5$. ■

Example 53: Find $\dim(S \cap T)$ where S and T are subspaces of the vector space $\mathbb{R}^4$ given by:

[2] **(2020.CC6.3ci)**

SOLUTION: উত্তরটা তুমি degree of freedom-এর কায়দায় চোখে দেখেই বুঝে যেতে পারো। এখানে কাজ হচ্ছে $\mathbb{R}^4$-এর মধ্যে, যার প্রতিটা সদস্যকেই দেখতে (x, y, z, w)-এর মত। চারটে অক্ষর, তাই চারটে degrees of freedom. যেই তুমি S-এর মধ্যে ঢুকবে, অমনি একটা শর্ত চাপবে, $2x + y + 3z + w = 0$. এ থেকে তুমি w-কে বাকিগুলো দিয়ে লিখে ফেলতে পারবে, $w = -2x - y - 3z$. সুতরাং S-এর সদস্যরা দেখতে হবে $(x, y, z, -2x - y - 3z)$-এর মত। তিনটে অক্ষর লাগছে, তাই $\dim(S)$ হল 3. এবার যখন $S \cap T$-এর মধ্যে ঢুকছি, তখন আরো একটা শর্ত চাপছে, $x + 2y + z + 3w = 0$, মানে $x + 2y + z + 3(-2x - y - 3z) = 0$, বা $-5x - y - 8z = 0$, যা থেকে আমরা y-কে বাকিগুলো দিয়ে লিখতে পারছি, $y = -5x - 8z$. সুতরাং $w = -2x - y - 3z = -2x + 5x + 8z - 3z = 3x + 5z$. সব মিলিয়ে বলতে পারি যে $S \cap T$-এর সদস্যরা দেখতে $(x, -5x - 8z, z, 3x + 5z)$-এর মতন। যেহেতু এখানে খালি দুটো অক্ষর লাগছে, তাই $\dim(S \cap T) = 2$. লক্ষ্য করো, এখানে শুরুতে ছিল চারটে অক্ষর, প্রতিটা শর্তের জন্য একটা করে অক্ষর কমল, কারণ শর্তদুটো linearly independent ছিল। এইভাবে চিন্তা করলে কোনো পরিশ্রম না করেই dimension-টা বুঝে ফেলা যায়। কিন্তু যদি সেটা লিখে প্রমাণ করতে হয়? তা হলেও এই একই কায়দায় এগোনো যায়--

Here

$$
\begin{aligned}
S\cap T &= \{(x,y,z,w)\in\mathbb{R}^4 \;:\; 2x+y+3z+w=0,\quad x+2y+z+3w=0\}
\end{aligned}
$$

এবার এক্ষুণি যে অক্ষর কমানোর ব্যাপারটা করলাম, সেটাই করব।

$$
\begin{aligned}
&= \cdots\\
&= \{(x,-5x-8z,z,3x+5z) \;:\; x,z\in\mathbb{R}\}\\
&= \{x(1,-5,0,3)+z(0,-8,1,5) \;:\; x,z\in\mathbb{R}\}.
\end{aligned}
$$

So $\{(1,-5,0,3),\ (0,-8,1,5)\}$ is a spanning set for $S\cap T$. Also these two vectors are not multiples of each other, and so linearly independent.

Thus $\{(1,-5,0,3),\ (0,-8,1,5)\}$ is a basis of $S\cap T$. Hence $\dim(S\cap T)=2$.

■

$\dim(S\cap T)$ সবচেয়ে কম হতে পারে 0 আর সবচেয়ে বেশী হতে পারে $\{\dim(S),\dim(T)\}$. যে কায়দায় উপরের উদাহরণটা করলাম, তা থেকেই বুঝতে পারছ যে, এ দুয়ের মধ্যবর্তী যেকোনো value-ই $\dim(S\cap T)$-র পক্ষে নেওয়া সম্ভব, অর্থাৎ যদি তোমাকে এমন যেকোনো $p,q,r\in\{0,1,2,...\}$ দিই, যাতে $r\leq p,q$ হয়, তবে অমনি তুমি এমন একটা vector space দিতে পারবে V, যার মধ্যে S,T দুটো subspace, এবং $\dim(S)=p$, $\dim(T)=q$ এবং $\dim(S\cap T)=r$ হয়।

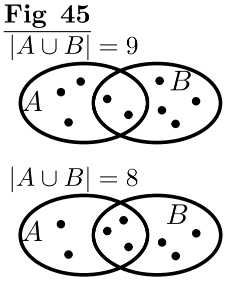
এটাকে একটা উপমা দিয়ে ভাবা যায়। যদি তোমাকে দুটো finite set দিতাম, যাদের সাইজ 5 আর 6, তবে তাদের intersection-এর সাইজ যেমন 0 থেকে 5 পর্যন্ত যা খুশি হতে পারে, এও তেমনি। এইভাবে set-এর সাইজের সঙ্গে vector space-এর dimension-এর উপমাটা আরো এগিয়ে নিয়ে যাওয়া যায়। যদি বলি দুটো set আছে, A আর B, যাদের সাইজ যথাক্রমে 5 এবং 6. তবে $A\cup B$-এর সাইজ কত হবে? এর উত্তর কিন্তু $5+6=11$ নাও হতে পারে (Fig 45)। উত্তরটা নির্ভর করছে ওদের মধ্যে কতটা intersection আছে তার উপর, কারণ যদি তুমি স্রেফ ওদের সাইজ দুটো যোগ করে দাও, তবে ওই intersection-টা দুইবার হিসেবে ঢুকে যাবে। তাই সঠিক উত্তরটা পাওয়ার জন্য intersection-এর সাইজটাকে ওই যোগফল থেকে একবার বাদ দিতে হবে। ফলে আসবে এই পরিচিত ফর্মুলাটা--

$$|A\cup B| = |A|+|B|-|A\cap B|.$$

Vector space-দের বেলায় অবশ্য S আর T দুটো vector space হলে $\dim(S\cup T)$ নিয়ে আলোচনা করা যায় না, কারণ $S\cup T$ একটা subspace নাও হতে পারে। কিন্তু একই যুক্তি খাটানো যায় $S+T$-এর ব্যাপারে। এবার সেটাই আলোচনা করব।

14.1.2 Sum

Fig 45



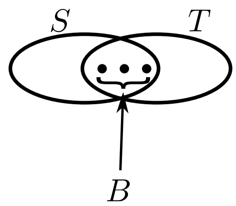


Fig 46

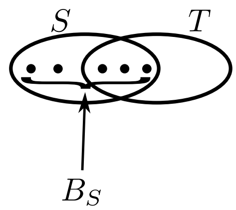


Fig 47

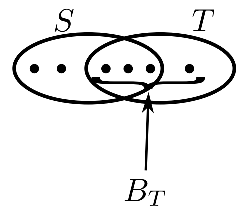


Fig 48

THEOREM

If S, T are two finite dimensional subspaces of a vector space, then

$$\dim(S+T) = \dim(S) + \dim(T) - \dim(S \cap T).$$

Proof: আমরা এক্ষুণি set-এর সাইজের যে উপমাটা দিলাম, সেটা ব্যবহার করেই প্রমাণটা হবে। আমরা শুরু করব $S \cap T$-এর একটা basis নিয়ে (ধরো B)। ছবি দেখতে চাইলে Fig 46 দেখে নাও। সেটাকে extend করব S-এর একটা basis-এ (ধরো B_S), (Fig 47) এবং আলাদা করে T-এর একটা basis-এ (ধরো B_T) (Fig 48)। এবার দেখব $B_S \cup B_T$ হবে $S+T$-এর একটা basis. ব্যস্, এবার $B_S \cup B_T$-এর সাইজ বার করব ওই "intersection-টা একবার বাদ দেওয়া"-র কায়দায়। তাতেই প্রমাণটা হয়ে যাবে।

> Step 1: Let $\{\vec{u}_1, ..., \vec{u}_p\}$ be a basis of $S \cap T$.
> We extend it to a basis $\{\vec{u}_1, ..., \vec{u}_p, \vec{s}_1, ..., \vec{s}_q\}$ of S.
> Separately, we extend it to a basis $\{\vec{u}_1, ..., \vec{u}_p, \vec{t}_1, ..., \vec{t}_r\}$ of T.
> Shall show that $\{\vec{u}_1, ..., \vec{u}_p, \vec{s}_1, ..., \vec{s}_q, \vec{t}_1, ..., \vec{t}_r\}$ is a basis of $S+T$.

এইটা আমরা এক্ষুণি দেখাব। কিন্তু তার আগে বলে নিই এটা করলেই কেন প্রমাণটা হয়ে যাবে।

> This will complete the proof, because $\dim(S \cap T) = p$, $\dim(S) = p+q$, $\dim(T) = p+r$ and $\dim(S+T) = p+q+r$. So
>
> $$\dim(S+T) = p+q+r = (p+q)+(p+r)-p = \dim(S)+\dim(T)-\dim(S\cap T).$$

এইবার $S+T$-এর basis দেখানো শুরু করব। প্রথমে দেখাব spanning set, তারপর দেখাব linearly independent.

> Step 2: Shall show that $\{\vec{u}_1, ..., \vec{u}_p, \vec{s}_1, ..., \vec{s}_q, \vec{t}_1, ..., \vec{t}_r\}$ spans $S+T$.
> Take any $\vec{v} \in S+T$.

$S+T$-এর প্রতিটা সদস্যই S-এর কোনো সদস্যর সঙ্গে T-এর কোনো সদস্য যোগ করে পাওয়া। তাই--

> Then $\vec{v} = \vec{s} + \vec{t}$ for some $\vec{s} \in S$ and $\vec{t} \in T$.

এবার আমরা $\vec{s}$-কে S-এর basis, এবং $\vec{t}$-কে T-এর basis দিয়ে লিখব।

Since $\{\vec{u}_1, ..., \vec{u}_p, \vec{s}_1, ..., \vec{s}_q\}$ is a basis of S, we must have

$$\vec{s} = \sum_i \alpha_i \vec{u}_i + \sum_j \alpha_{p+j} \vec{s}_j$$

for some $\alpha_1, ..., \alpha_{p+q}$.
Similarly, since $\{\vec{u}_1, ..., \vec{u}_p, \vec{t}_1, ..., \vec{t}_r\}$ is a basis of T, we must have

$$\vec{t} = \sum_i \beta_i \vec{u}_i + \sum_k \beta_{p+k} \vec{t}_k$$

for some $\beta_1, ..., \beta_{p+r}$.
So

$$\vec{v} = \vec{s} + \vec{t} = \sum_i (\alpha_i + \beta_i) \vec{u}_i + \sum_j \alpha_{p+j} \vec{s}_j + \sum_k \alpha_{p+k} \vec{t}_k.$$

তার মানে যেকোনো $\vec{v} \in S + T$-কেই $\vec{u}_i$, $\vec{s}_j$ আর $\vec{t}_k$-দের linear combination হিসেবে লেখা যাচ্ছে।

Hence $\{\vec{u}_1, ..., \vec{u}_p, \vec{s}_1, ..., \vec{s}_q, \vec{t}_1, ..., \vec{t}_r\}$ spans $S + T$, as required.

এবার linearly independent দেখাব--

Step 3: Shall show that $\{\vec{u}_1, ..., \vec{u}_p, \vec{s}_1, ..., \vec{s}_q, \vec{t}_1, ..., \vec{t}_r\}$ is linearly independent.

যথারীতি আমরা একটা linear combination-কে $\vec{0}$ ধরে নিয়ে দেখাব যে, সবগুলো coefficient-ই শূন্য হতে বাধ্য।

Let $\sum_i a_i \vec{u}_i + \sum_j b_j \vec{s}_j + \sum_k c_k \vec{t}_k = \vec{0}$.
Shall show $\forall i \quad a_i = 0$, $\forall j \quad b_j = 0$ and $\forall k \quad c_k = 0$.
Then $\sum_i a_i \vec{u}_i + \sum_j b_j \vec{s}_j = -\sum_k c_k \vec{t}_k$.
The LHS is in S and the RHS is in T. So both sides must be in $S \cap T$.

এই লাইনটা ভালো করে বুঝে নাও। পুরো প্রমাণটার মধ্যে এইটাই হল প্যাঁচ।

Hence we can write this as $\sum_i d_i \vec{u}_i$ for some d_i's.

কারণ $\vec{u}_i$-গুলো হল $S \cap T$-এর basis.

So $-\sum_k c_k \vec{t}_k = \sum_i d_i \vec{u}_i$,
or $\sum_k c_k \vec{t}_k + \sum_i d_i \vec{u}_i = \vec{0}$.
But, since $\{\vec{u}_1, ..., \vec{u}_p, \vec{t}_1, ..., \vec{t}_r\}$ is linearly independent, we have

$$\forall k \quad c_k = 0 \text{ and } \forall i \quad d_i = 0.$$

এর মধ্যে $d_i = 0$ নিয়ে আমাদের মাথাব্যথা নেই। আমাদের দরকার ছিল a_i, b_j, c_k-দের শূন্য দেখানো। তার মধ্যে c_k-দের ধরাশায়ী করা গেল। আমরা শুরু করেছিলাম $\sum_i a_i \vec{u}_i + \sum_j b_j \vec{s}_j + \sum_k c_k \vec{t}_k = \vec{0}$ নিয়ে। সেখান থেকে c_k-রা বাদ গেলে পড়ে থাকে--

> Hence $\sum_i a_i\vec{u}_i + \sum_j b_j\vec{s}_j = -\vec{0}$.
> Now, since $\{\vec{u}_1, ..., \vec{u}_p, \vec{s}_1, ..., \vec{s}_q\}$ is linearly independent,
> we have $\forall i \quad a_i = 0$ and $\forall j \quad b_j = 0$, as required.

[Q.E.D]

Exercise 52: Show that $U = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \ : \ a+b=0 \right\}$, $W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \ : \ c+d=0 \right\}$ are the subspaces of $\mathbb{R}_{2\times 2}$. Find the dimensions of U, $U\cap W$ and $U+W$.[2+1+1+1] **(2020.CC6.3d)**
HINT: এখানে degrees of freedom দিয়ে চিন্তা করো। সেভাবে U, W এবং $U\cap W$-এর dimension বেরিয়ে যাবে। তারপর সেখান থেকে $\dim(U+W)$ বার করতে কতক্ষণ? ■

এবার এক ধরণের অংক দেখব যেখানে finite set-এর সাইজের সঙ্গে উপমাটা মনে রাখলে সুবিধা হবে।

Example 54: ধরো একটা set আছে, যার সাইজ 6. এর মধ্যে দুটো subset বানাতে হবে, যাদের একটার সাইজ 5, অন্যটার 4. তবে ওদের intersection-এর সাইজ কি 2 হতে পারে?
SOLUTION: যেহেতু $2 \leq 5, 6$, তাই ফস্ করে মনে হতে পারে যে, intersection-এর সাইজ 2 হওয়া অসম্ভব নয়। কিন্তু মনে রেখো পুরো জিনিসটা ফিট করতে হবে একটা 6 সাইজের set-এর মধ্যে। তাই গাদাগাদি না করলে চলবে না। যদি subset দুটোকে A আর B বলি, তবে $|A\cup B| \leq 6$ হতে হবে, মানে-- $|A|+|B|-|A\cap B| \leq 6$ দরকার, অর্থাৎ $|A\cap B| \geq |A|+|B|-6 = 5+4-6 = 3$ হতে হবে। একটু চিন্তা করলেই বুঝবে যে, intersection-এর সাইজ 3 হতেই পারে। ■

একই যুক্তিতে নীচের অংকটা করো।

Exercise 53: If S and T are two subspaces of $\mathbb{R}^{10}$ with $\dim(S) = 7$ and $\dim(T) = 9$, then what is the least possible dimension of $S\cap T$? ■

Set-এর সাইজের সঙ্গে vector space-এর dimension-এর উপমাটা এখানেই শেষ নয়। ধরো U হল যাকে বলে গিয়ে universal set, আর তার মধ্যে রয়েছে A. যদি $|U| = n$ হয়, তবে $|A^c| = n - |A|$ হয়। একই কথা খাটে subspace-দের বেলাতেও। আমরা দেখেছি যে যদি S একটা subspace হয়, তবে S^c একটা subspace হতে পারে না। কিন্তু complement-এর ধারণাটা subspace-দের বেলাতেও আছে, দুই রূপে--এক, direct complement আর দুই, quotient space হিসেবে। দুইক্ষেত্রেই dimension-টা একই হয়। প্রথমে আমরা direct complement নিয়ে আলোচনা করব।

14.1.3 Direct complement

> THEOREM
>
> Let V be a finite dimensional vector space. Let S be a subspace of V. Then for any direct complement T of S in V, we have $\dim(T) = \dim(V) - \dim(S)$.

যদি তুমি basis extension দিয়ে চিন্তা করো, তবে theorem-এর প্রথম অংশটা বলাই বাহুল্য। কারণ S-এর প্রতিটা direct complement-ই আসে এইভাবে-- S-এর একটা basis নিয়ে শুরু করতে হয়, ধরো B_S. সেটাকে extend করে পুরো

V-এর একটা basis বানাতে হয়, B_V. এর জন্য বাড়তি যে কটা vector লাগল, মানে $B_V \setminus B_S$-এর সদস্যরা, তারা হয় direct complement-টার basis.
এখানে $|B_S| = \dim(S)$ আর $|B_V| = \dim(V)$.
তাই $|dim(T)| = |B_V \setminus B_S| = |B_V| - |B_S| = \dim(V) - \dim(S)$.

Exercise 54: যদি $S = \{(x, y, z) \ : \ x + y + z = 0\}$ হয়, তবে $\mathbb{R}^3$-এর মধ্যে ওর যেকোনো direct complement-এর dimension কত হবে? ■

Exercise 55: যদি $\mathbb{R}^4$-এর একটা subspace হয় S, যার একটা direct complement হল

$$T = \{(x, y, z, w) \ : \ x + y + z = 0, \ y + z + w = 0\},$$

তবে $\dim(S) =$? ■

এবারের অংকটা একটু কঠিন।

Exercise 56: যদি $\mathbb{R}^n$-এর মধ্যে S একটা 1-dimensional subspace হয়, এবং U আর V ওর দুটো direct complement হয় (যেখানে $U \neq V$), তবে $\dim(U \cap V)$ কত হবে? ■

14.1.4 Quotient

THEOREM

Let V be a finite dimensional vector space. Let W be a subspace of V. Then

$$\dim(V/W) = \dim(V) - \dim(W).$$

প্রমাণটা চেয়েছে নীচের অংকে। সেটায় হাত দেওয়ার আগে quotient space-এর যোগ আর গুণ একটু মনে করিয়ে দিই--

- V/W-এর মধ্যে $\vec{u} + W$ আর $\vec{v} + W$-কে যোগ করলে হয় $(\vec{u} + \vec{v}) + W$.
- একইভাবে $\vec{u} + W$-কে α দিয়ে গুণ করলে হয় $\alpha\vec{u} + W$.
- যদি $\vec{v} + W = W$ হয়, তবে $\vec{v} \in W$ হবে। বিপরীতপক্ষে, যদি $\vec{v} \in W$ হয়, তবে $\vec{v} + W = W$ হতে বাধ্য।
- আরেকটা কথা, V/W-র মধ্যে zero vector-টা হল W.

Example 55: Let V be a finite-dimensional vector space over a field F and W be a subspace of V. Show that $\dim(V/W) = \dim(V) - \dim(W)$.[5] **(2021.CC6.3b)**
SOLUTION: প্রমাণটা পড়ার সময়ে একটা উদাহরণ মাথায় রাখলে সুবিধা হবে। ধরো V হল $\mathbb{R}^3$. তার মধ্যে W হল x-axis-টা। তাহলে V/W হল Fig 49-এর মত দেখতে। আমরা এর একটা basis বানাব। এর প্রত্যেকটা সদস্য হল $\vec{v} + W$ চেহারার। আমরা এমন কিছু $\vec{v}_1, ..., \vec{v}_n \in V$ বার করব, যাতে $\vec{v}_1 + W, ..., \vec{v}_n + W$-রা মিলে V/W-এর একটা basis তৈরী করে।

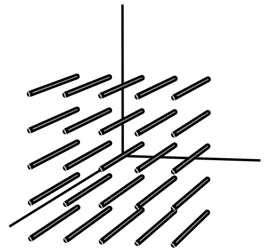


Fig 49

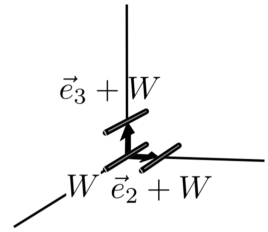


Fig 50

> Let $\{\vec{w}_1, ..., \vec{w}_m\}$ be a basis of W. Extend it to a basis $\{\vec{w}_1, ..., \vec{w}_m, \vec{v}_1, ..., \vec{v}_n\}$ of V.

আমাদের $\mathbb{R}^3$-র উদাহরণে আমরা W-এর basis নেব $\{\vec{e}_1\}$. সেটাকে extend করব $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$-তে। এর ফলে বাড়তি এল $\vec{e}_2$ আর $\vec{e}_3$. আমাদের দাবী হল $\{\vec{e}_2 + W, \vec{e}_3 + W\}$ হবে V/W-এর একটা basis. আমরা Fig 50-এ $\vec{e}_2 + W$ আর $\vec{e}_3 + W$-কে দেখিয়েছি। আশা করি অনুভব করতে পারছ কেন এরা basis হচ্ছে--বার বার করে $\vec{e}_2 + W$ ব্যবহার করে মেঝেটা বিছিয়ে ফেলতে পারবে, আর $\vec{e}_3 + W$ ব্যবহার করে উপরনীচে ওঠানামা করতে পারবে। এবার প্রমাণটার দিকে তাকানো যাক।

> Then $n = \dim(V) - \dim(W)$.
>
> Shall show that $\{\vec{v}_1 + W, ..., \vec{v}_n + W\}$ is a basis of V/W to complete the proof.

যথারীতি দুটো অংশ, spanning আর linearly independent.

> <u>Spanning:</u> Take any $\vec{v} + W$ where $\vec{v} \in V$.
>
> Then $\vec{v} = \sum a_i \vec{w}_i + \sum b_j \vec{v}_j$.
>
> So $\vec{v} + W = (\sum a_i \vec{w}_i + \sum b_j \vec{v}_j) + W = \sum b_j \vec{v}_j + W$, since $\forall i \quad \vec{w}_i \in W$.

এইটা বুঝলে তো? যদি $\vec{w} \in W$ হয়, তবে $\vec{w} + W = W$ হবে।

> Thus $\vec{v} + W = \sum b_j (\vec{v}_j + W)$, as required.

এবার linearly independent দেখাব। তার জন্য একটা linear combination-কে zero vector-এর সমান করে দেখাব যে প্রতিটা coefficient-ই আলাদা করে শূন্য হতে বাধ্য। মনে রেখো এখানে W-টাই zero vector-এর ভূমিকা পালন করছে।

> <u>Linearly independent:</u> Let $\sum c_i (\vec{v}_i + W) = W$.
>
> Then $(\sum c_i \vec{v}_i) + W = W$.
>
> $\therefore \sum c_i \vec{v}_i \in W$.
>
> So $\sum c_i \vec{v}_i = \sum d_j \vec{w}_j$ for some $b_i \in F$.

কারণ $\vec{w}_j$-রা হল W-এর basis.

> But this means $\sum c_i\vec{v}_i - \sum d_j\vec{w}_j = \vec{0}$.
> As $\{\vec{w}_1, ..., \vec{w}_m, \vec{v}_1, ..., \vec{v}_n\}$ is linearly independent, hence this implies $\forall i \quad c_i = 0$ and $\forall j \quad d_j = 0$.
> This completes the proof.

এখানে অবশ্য d_j-গুলো শূন্য হওয়া নিয়ে আমাদের মাথাব্যথা ছিল না, c_i-গুলোকে যে শূন্য দেখানো গেছে, এতেই আমরা খুশি! ■

DAY 15
Reduced row-echelon form (RREF)

> সেই গন্ধবিকট তেলের কথা শোন নি? সেই তেলের আশ্চর্য গুণ!...সে তেল খেলে পরে পিলের ওষুধ, মাখলে পরে ঘায়ের মলম, আর গোঁফে লাগালে দেড় দিনে আধ হাত লম্বা গোঁফ বেরোয়।
>
> --সুকুমার রায়

আজকে যেটা আলোচনা করব, সেটা সিলেবাসে সরাসরি নেই, কিন্তু যেটা জানা দরকারী। বিশেষ করে যারা বিভিন্ন প্রতিযোগিতামূলক পরীক্ষায় আগ্রহী, সেখানে এটা কাজে লাগবে, কারণ কোনো MCQ-এর উত্তর চট্ করে রাফে বার করতে RREF কাজে দিতে পারে। বিশেষতঃ আজকের এই কম্পিউটারের যুগে linear algebra-র বিভিন্ন প্রয়োগ প্রত্যক্ষ বা পরোক্ষভাবে এর উপরে দাঁড়িয়ে আছে। আমরা এখনো পর্যন্ত অনেক কিছু আলোচনা করেছি, যেমন span, linear independence, subspace নিয়ে বিভিন্ন কাজ, basis-কে extend করা, যেগুলো খুব ছোটোছোটো উদাহরণে হাতে কষে ফেলা সহজ, কিন্তু একটু লম্বা উদাহরণ হলেই মাথা গুলিয়ে যায়। এই সব কাজ করার অমোঘ অস্ত্র হল **reduced row echelon form (RREF)**, যেটা আমরা এবার শিখব। এই জিনিসটার অনেকরকম গুণ। ধরো কিছু vector-এর একটা তালিকা দেওয়া আছে--

$$\vec{v}_1, ..., \vec{v}_k.$$

তাহলে RREF-এর সাহায্যে আমরা কত কিছু করতে পারি তার ফিরিস্তি শোনো--

- আমরা বলে দিতে পারব এরা linearly independent কিনা।
- এরা মিলে পুরো vector space-টাকে generate করে কিনা।
- যদি কোনো vector দেওয়া থাকে $\vec{v}$, তবে $\vec{v} \in \text{span}\{\vec{v}_1, ..., \vec{v}_k\}$ হবে কিনা।
- চাইলে $\text{span}\{\vec{v}_1, ..., \vec{v}_k\}$-এর একটা basis-ও বার করে দেওয়া যাবে।
- কোনো linearly independent set-কে basis-এ extend-ও করা যাবে RREF দিয়ে।

এছাড়াও আরো অনেক কিছু জানা যায় RREF থেকে, যেগুলো এ বইতে যথাস্থানে বলব। সুতরাং এক কথায় বলতে গেলে RREF হল সাক্ষাৎ কল্পতরু, এর থেকে এক ধাক্কায় এক রাশ দরকারী তথ্য বেরিয়ে আসে। ব্যাপারটা এবার আমরা কিছু উদাহরণের মাধ্যমে ধাপে ধাপে বুঝব।

Example 56: বলতে হবে এই set-টা linearly independent কিনা--

$$\{(1, 7, 3, 4), (2, 14, 5, -1), (4, 3, 2, 5)\}.$$

SOLUTION: এখানে দেখতে হবে যদি

$$a(1,7,3,4)+b(2,14,5,-1)+c(4,3,2,5)=(0,0,0,0)$$

হয়, তবে $a=b=c=0$ হতে বাধ্য কিনা। কাজটা স্রেফ চোখে দেখে করা শক্ত। Equation solve করতে হবে। আমরা অবশ্য solve করার মধ্যে যাচ্ছি না, কাজটা যে চোখে দেখেই করে ফেলা যায় না, খালি সেটুকু বোঝানোই এখানে আমাদের উদ্দেশ্য। ■

বুঝতেই পারছ যদি আরও বেশী সংখ্যক vector থাকত, তবে কাজটা উত্তরোত্তর কঠিন হয়ে উঠত। তার কারণ-- কিছু vector দেওয়া থাকলে তাদের linear combination করে কী কী vector পাওয়া যায়, সেটা চট্ করে বোঝা যায় না। তবে সৌভাগ্যক্রমে এর একটা ব্যতিক্রম আছে, যেখানে উত্তরটা চোখে দেখেই ধাঁ করে বলে দেওয়া যায়। সেটাই দেখিয়েছি নীচের উদাহরণে। এর মূলমন্ত্র একটাই, 0-কে তুমি যা দিয়েই গুণ কর না কেন উত্তর সর্বদা 0-ই হয়।

Example 57: এই set-টা linearly independent কিনা বল--

$$\{(1,2,3,4,6),(0,-2,3,45.6,5),(0,0,67,6,23),(0,0,0,4,5)\}.$$

SOLUTION: Vector চারটের নাম দিয়ে নিই, $\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4$. এবার শেষের থেকে দেখা শুরু করো। $\vec{v}_3\in\text{span}\{\vec{v}_4\}$ কী হতে পারে? না, কারণ তিন নম্বর entry-টা $\vec{v}_4=(0,0,\mathbf{0},4,5)$-এর বেলায় 0, তাকে যা দিয়েই গুণ কর না কেন, $\vec{v}_3=(0,0,\mathbf{67},6,23)$-এর তিন নম্বর স্থানের $67(\neq 0)$ কোনো দিনই পাবে না।
এবার বল $\vec{v}_2\in\text{span}\{\vec{v}_3,\vec{v}_4\}$ হতে পারে কি না। অসম্ভব, $\vec{v}_3$ আর $\vec{v}_4$ দুজনেরই দ্বিতীয় স্থানে 0 আছে। সুতরাং যতই ওদের linear combination বানাও না কেন, দ্বিতীয় স্থানের 0-টা ঘুচবে না। কিন্তু $\vec{v}_2$-এর দ্বিতীয় স্থানে আছে $-2(\neq 0)$. ঠিক একইভাবে প্রথম স্থানের দিকে তাকালেই বুঝবে যে $\vec{v}_1\notin\text{span}\{\vec{v}_2,\vec{v}_3,\vec{v}_4\}$.
সুতরাং চোখে দেখেই বলে দিতে পারলাম যে, এই set-টা linearly independent. ■

এই কায়দাটা খুবই কাজের, কিন্তু দুঃখের কথা হল জুৎসই মত জায়গায় অনেকগুলো 0 না থাকলে এই কায়দাটা কাজ করে না। আসলে 0 (এবং 1-এরও) অন্যতম সুবিধা হল ওদের দিয়ে অনায়াসে মনে মনেই গুণ করে ফেলা যায়। Reduced row echelon form (RREF) ব্যাপারটার ভূমিকা এখানেই। RREF মানে আর কিছুই নয়-- "জুৎসই জায়গায় কিছু 0 আর 1 থাকা"। বিশদ করেই বলি।

15.1 সংজ্ঞা

আমাদের কাছে কিছু vector-এর একটা তালিকা আছে। প্রতিটা vector এককেটা tuple-এর মত কল্পনা করো[4], এবং তাদেরকে row-এর মত একটার উপর একটা শুইয়ে একটা matrix বানাও। এই matrix-টা যদি তিনটে শর্ত পালন করে, তবে বলব যে সেটা RREF রূপে আছে--

1. প্রথম শর্তটা nonzero row-দের নিয়ে, মানে যেসব row-তে অন্ততঃ একটাও nonzero entry আছে। এদের মধ্যে সবচেয়ে বাঁদিকে যে nonzero entry-টা আছে সেটাকে 1 হতেই হবে। ওকে বলে সেই row-এর **leading** 1. কোনো leading 1-এর উপরে বা নীচে 0 ছাড়া আর কিছু না থাকতে পারবে না।

যেমন $\begin{bmatrix}0&\mathbf{1}&2&0\\0&0&0&\mathbf{1}\\0&0&0&0\end{bmatrix}$ রয়েছে RREF রূপে। এখানে leading 1-দের মোটা করে দেখিয়েছি। কিন্তু $\begin{bmatrix}0&1&2&0\\0&0&0&\mathbf{2}\\0&0&0&0\end{bmatrix}$ RREF রূপে নেই। একইভাবে $\begin{bmatrix}0&1&2&\mathbf{3}\\0&0&0&1\\0&0&0&0\end{bmatrix}$-ও RREF রূপে নেই।

[4] যদি vector-গুলো tuple-জাতীয় না হয়, তবে কোনো basis-এর সাপেক্ষে ওদেরকে tuple-এ পরিণত করে নিতে হবে। যেমন $1+2x+3x^2$-কে $\{1,x,x^2\}$-এর সাপেক্ষে $(1,2,3)$ লেখা যায়।

2. matrix-টা খুব বড় হলে nonzero row-দের চট করে চোখে দেখে খুঁজে পাওয়া শক্ত হতে পারে। তাই একটু সাজিয়ে লিখলে সুবিধা হয়--যাবতীয় nonzero row-রা থাকবে উপরে, এবং zero row-রা (মানে যেসব row-এর সবগুলো entry-ই 0) তারা থাকবে সবার নীচে। এরকম row থাকতেই হবে এমন বলছি না, কিন্তু যদি থাকে তবে তাদের নীচে যেন আর কোনো nonzero row না থাকে।

   যেমন $\begin{bmatrix} 0 & 1 & 2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0 & 0 & 0 & 1 \end{bmatrix}$-এর মধ্যে zero row-টা একটা nonzero row-এর উপরে উঠে গেছে, তাই RREF হচ্ছে না।

3. যদি matrix-টা খুব বড় হয়, তবে চোখে দেখে leading 1-দের খুঁজে পেতে সময় লেগে যেতে পারে, তাই nonzero row-গুলোকে নিজেদের মধ্যে একটু গুছিয়ে লিখলে সুবিধা হয়। সেই প্রয়োজন থেকেই আসে তৃতীয় শর্তটা--প্রথম nonzero row-এর leading 1-টা যে column-এ আছে, দ্বিতীয় nonzero row-এর বেলায় যেন তার ডানদিকে থাকে, তৃতীয় nonzero row-এর বেলায় থাকে তারও ডানদিকে, এরকম। অর্থাৎ যত নীচে নামবে, leading 1-রাও ততই ডানদিকে সরবে।

   যেমন $\begin{bmatrix} 0 & 0 & 0 & \mathbf{1} \\ 0 & \mathbf{1} & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ কিন্তু RREF-এ নেই।

এখানে প্রথম শর্তটাই এখানে সবচেয়ে গুরুত্বপূর্ণ, বাকি দুটো শর্ত হল খালি row-গুলোকে সুন্দর করে সাজিয়ে লেখার জন্য। RREF-এর সুবিধা দুটো--

- এক, RREF-এর nonzero row-গুলো linearly independent হয়। এটা কেন হয়, সেটা বোঝা শক্ত নয়, কারণ leading 1-দের উপরে নীচে 0 থাকে।
- দুই, যদি দুটো সমান দৈর্ঘ্যের তালিকা থাকে, যারা একই subspace-কে span করে, এবং দুজনেই RREF-এ আছে, তবে ওরা একই হতে বাধ্য। এটা প্রমাণ করা একটু শক্ত। আমরা এখানে প্রমাণটার মধ্যে যাব না। তবে একটা উপমা দিয়ে ভাবতে পারো। Vector-এর তালিকাটা যেন একটা ভগ্নাংশ, আর span হল যেন ভগ্নাংশটার value. দুটো বিভিন্ন তালিকার যেমন একই span হতে পারে, ঠিক তেমনি $\frac{2}{4}$ আর $\frac{3}{6}$ ভগ্নাংশদুটোকে দেখতে আলাদা হলেও ওদের একই value. এই উপমায় RREF ঠিক যেন ভগ্নাংশের সহজতম রূপ, মানে যেখানে আর কাটাকাটি করা যায় না, এবং denominator হয় > 0. যদি দুটো ভগ্নাংশের value যদি সমান হয়, এবং দুজনেই সহজতম রূপে থাকে, তবে ওরা সমান হতে বাধ্য। একইভাবে দুটো RREF তালিকার span সমান হলে ওরা একই হতে বাধ্য।

15.2 Gauss Jordan elimination

যদি কিছু vector-এর একটা তালিকা দেওয়া থাকে, তবে আমরা **Gauss-Jordan elimination** নামে একটা প্রক্রিয়া করে ওকে একই দৈর্ঘ্যের এমন একটা তালিকায় পরিণত করে ফেলতে পারি, যার span মূল তালিকাটার span-এর সমান হয়। ভগ্নাংশের উপমা দিয়ে ভাবলে এটা একটা ভগ্নাংশকে কাটাকাটি করে সহজতম রূপে লেখার মত। কাটাকাটি করলে যেমন ভগ্নাংশের value-টা বদলায় না, তেমনি Gauss-Jordan elimination করলে তালিকাটার span বদলায় না।
কোনো ভগ্নাংশ হাতে পেলেই যেমন আমরা ওটাকে কাটাকাটি করে সহজ করে নিতে ভালোবাসি, তেমনি কিছু vector-এর একটা তালিকা পেলেই ওটাকে Gauss-Jordan elimination দিয়ে RREF-এ পরিণত করে নিলে অনেক সুবিধা হয়।
এবার আমরা Gauss-Jordan elimination প্রক্রিয়াটা শিখব। এর মূল উপাদান হল তিনটে কাজ, যাদের বলে এককেটা **basic operation.**

15.2.1 তিনটে basic operation

ধরো S হল কিছু vector-এর একটা তালিকা, $\vec{v}_1, ..., \vec{v}_n$. এই তালিকার উপর তিনরকম operation শিখব, যার সাহায্যে নতুন একটা তালিকা তৈরী হবে। প্রতি ক্ষেত্রেই তালিকাটার দৈর্ঘ্য, span এবং linear independence অপরিবর্তিত থাকবে।

অর্থাৎ নতুন তালিকাটাকে যদি T বলি তবে S-এ যতগুলো vector ছিল, T-তেই তাই থাকবে, $\text{span}(S) = \text{span}(T)$ হবে এবং S যদি linearly (in)dependent হয়, তবে T-ও তাই হবে,
তিনটে operation-কে আমরা নাম দিয়েছি swap, scale এবং add--

- Swap: তালিকার যেকোনো দুটো vector-এর স্থান বিনিময় করে দেওয়া। যেমন যদি S হয় $\vec{v}_1, \vec{v}_2, \vec{v}_3$, তবে শেষের দুটোর স্থান বিনিময় করে দিলে পাব এই তালিকাটা $\vec{v}_1, \vec{v}_3, \vec{v}_2$. এইটাই আমাদের T.
- Scale: তালিকার কোনো একটা vector-কে কোনো nonzero scalar দিয়ে গুণ করে দেওয়া। যেমন, যদি শুরুতে তালিকাটা হয় $\vec{v}_1, \vec{v}_2, \vec{v}_3$, আর তুমি $\vec{v}_3$-কে 100 দিয়ে গুণ করো, তবে নতুন তালিকাটা হবে $\vec{v}_1, \vec{v}_2, 100\vec{v}_3$.
- Add: S থেকে যেকোনো একটা $\vec{v}_i$ নাও, এবং যেকোনো একটা $c \in F$ নাও[5]। তাহলে $c\vec{v}_i$ একটা vector হল, এইটাকে তুমি S-এর অন্য যেকোনো vector-এর সাথে যোগ করে দাও। এর ফলে যে নতুন তালিকাটা পাবে সেটাই হবে আমাদের T. যেমন ধরো, যদি S-এর vector-গুলো হয়

  $$\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$$

  আর তুমি $2\vec{v}_3$ যোগ করো $\vec{v}_2$-এর সাথে তবে তোমার T হবে

  $$\vec{v}_1, \vec{v}_2 + 2\vec{v}_3, \vec{v}_3, \vec{v}_4.$$

  লক্ষ্য করো, $\vec{v}_3$ কিন্তু অপরিবর্তিত রইল।

এইবার এই তিনটে operation দিয়ে কী করে RREF-এ পৌঁছতে হয়, সে কথা বলি।

15.2.2 Algorithm

> *ভাত মাখোরে, গ্রাস তোলরে, মুখের মধ্যে ঢোকাওরে, চিবোওরে, গেলোরে--তারপরে জল খাওরে, আঁচাওরে, হাত মুখ মোছরে! কতরকম কাণ্ড!*
>
> --সুকুমার রায়

Gauss-Jordan elimination হল একটা algorithm, অর্থাৎ ধাপে ধাপে কোনো কাজ করার কায়দা। ছোটোবেলায় আমরা LCM (লসাগু) আর HCF (গসাগু) বের করার যে কায়দা শিখি, সেগুলোও একেকটা algorithm. কোনো algorithm-কে বর্ণনা করে বোঝানোর একটাই সমস্যা। যে কাজটা করতে অতি সহজ, সেটাও বর্ণনা করতে গেলে ভয়ংকর রকমের কঠিন শোনায়। সুকুমার রায়ের একটা হাসির গল্পে একটা লোক ভাত খাওয়ার পদ্ধতির এমন বর্ণনা দিয়েছিল যে, শুনলে মনে হবে অমন কঠিন কাজ দুনিয়ায় আর দুটো নেই! তার চেয়ে বরং algorithm শেখার ভালো পথ হল কিছু উদাহরণ দেখে নেওয়া, যাতে কত রকম সম্ভাবনা আসতে পারে সবগুলোর সঙ্গে পরিচয় হয়ে যায়। এখানে আমাদের হাতে সময় কম। তাই এমন একটা উদাহরণ বানিয়েছি, যাতে সেই এক উদাহরণেই সবরকম সম্ভাবনার সঙ্গে পরিচয় হয়ে যাবে।

Example 58: ধরো তোমাকে এই তালিকাটা দিলাম--

$$(2, 4, 10, 32, -22), \quad (1, 2, 5, 17, -12), \quad (3, 8, 19, 61, -42), \quad (0, 1, 2, 5, -3).$$

প্রথমে এদেরকে একটার উপর একটা লিখে একটা matrix বানিয়ে নাও--

$$\begin{bmatrix} 2 & 4 & 10 & 32 & -22 \\ 1 & 2 & 5 & 17 & -12 \\ 3 & 8 & 19 & 61 & -42 \\ 0 & 1 & 2 & 5 & -3 \end{bmatrix}$$

[5] F হল আমাদের field-টা।

এবার এর উপরে Gauss-Jordan elimination লাগাও।
SOLUTION: প্রথমে শুরু করো $(1,1)$-th entry-টা থেকে। এই অবস্থানটাকে বলব এই ধাপে আমাদের **pivot**. খেয়াল রাখার সুবিধার জন্য এটাকে বাক্স এঁকে বুঝিয়েছি--

$$\begin{bmatrix} \boxed{2} & 4 & 10 & 32 & -22 \\ 1 & 2 & 5 & 17 & -12 \\ 3 & 8 & 19 & 61 & -42 \\ 0 & 1 & 2 & 5 & -3 \end{bmatrix}$$

আমাদের উদ্দেশ্য থাকে pivot-এর জায়গায় একটা leading 1 আনা। তার জন্য প্রথম দরকার ওখানে একটা nonzero কিছু থাকা। যদি ওখানে 0 থাকে, তবে আমরা ওখানে nonzero কিছু আনার চেষ্টা করি। যাই হোক, এখানে pivot-এর জায়গায় nonzero সংখ্যাই আছে, 2. এইটা দিয়ে প্রথম row-টাকে ভাগ করে ওখানে 1 করে নাও। তবে একটা leading 1 পেয়ে গেলে। সেই যে তিনটে basic operation-এর কথা বলেছিলাম, তার মধ্যে এখানে আমরা scale লাগালাম।

$$\xrightarrow{newR_1 = R_1/2} \begin{bmatrix} \boxed{1} & 2 & 5 & 16 & -11 \\ & & & & \\ & & & & \\ & & & & \end{bmatrix}$$

বাকি row-গুলো অপরিবর্তিত আছে, বলে আর নতুন করে লিখিনি। তীর চিহ্নর উপরে ওই যে $newR_1 = R_1/2$ লিখেছি, ওর মানে হল আগের matrix-টার প্রথম row-কে 2 দিয়ে ভাগ করে নতুন matrix-টার প্রথম row পাওয়া গেছে।
এই যে নতুন row-টা পেলে এর উপযুক্ত multiple বিয়োগ করে দাও অন্য row-গুলো থেকে, যাতে ওই leading 1-টার উপরে[6] আর নীচে সব entry শূন্য হয়ে যায়। Basic operation দিয়ে ভাবলে, এখানে আমরা দুবার add করব।

$$\xrightarrow[newR_3 = R_3 - 3R_1]{newR_2 = R_2 - R_1} \begin{bmatrix} 1 & 2 & 5 & 16 & -11 \\ 0 & \boxed{0} & 0 & 1 & -1 \\ 0 & 2 & 4 & 13 & -9 \\ 0 & 1 & 2 & 5 & -3 \end{bmatrix}$$

আমরা সফলভাবে প্রথম ধাপটা শেষ করলাম। যখনই সফল হব, তখনই pivot-টাকে নীচে ডানদিকে একঘর সরিয়ে নেব। তাই pivot-টা এখন চলে এসেছে $(2,2)$-th entry-তে। সমস্যা হল এখানে একটা 0 আছে। তাই leading 1 বানাতে পারছি না। যখনই এরকম সমস্যা হবে, তখনই তাই আমরা pivot-এর column বরাবর সোজা নীচের দিকে খুঁজব কোথাও কোনো nonzero entry আছে কিনা। এখানে নীচের দুটো entry-ই nonzero. এদের যেকোনো একটা row-এর সাথে এখনকার দ্বিতীয় row-টার স্থানবিনিময় করব, যাতে pivot-এর জায়গায় nonzero entry আসে। Basic operation দিয়ে ভাবলে, এটা একটা swap.

$$\xrightarrow{\text{Swap } R_2, R_3} \begin{bmatrix} 1 & 2 & 5 & 16 & -11 \\ 0 & \boxed{2} & 4 & 13 & -9 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 2 & 5 & -3 \end{bmatrix}$$

এখানে আমরা তৃতীয় row-এর সঙ্গে স্থানবিনিময় করলাম, চাইলে চতুর্থ row-এর সঙ্গেও করতে পারতাম। লক্ষ্য করো, pivot-এর অবস্থান কিন্তু বদলালো না। এখন সেই অবস্থানে nonzero সংখ্যা এসে গেছে, 2. আগেরই মত ফের এই nonzero entry দিয়ে ভাগ করে leading 1 বানাব--

[6]প্রথম ধাপে যেহেতু সবচেয়ে উপরের row-তেই আছি, তাই এর উপরে আর কোনো entry নেই।

$$\xrightarrow{newR_2 = R_2/2} \begin{bmatrix} 0 & \boxed{1} & 2 & \frac{13}{2} & -\frac{9}{2} \\ & & & & \end{bmatrix}$$

অপরিবর্তিত বাকি row-গুলোকে আর কষ্ট করে লিখি নি। এবার আবার leading 1-এর উপরে নীচের সবাইকে শূন্য বানানোর পালা--

$$\xrightarrow[newR_4 = R_4 - R_2]{newR_1 = R_1 - 2R_2} \begin{bmatrix} 1 & 0 & 1 & 3 & -2 \\ 0 & 1 & 2 & \frac{13}{2} & -\frac{9}{2} \\ 0 & 0 & \boxed{0} & 1 & -1 \\ 0 & 0 & 0 & -\frac{3}{2} & \frac{3}{2} \end{bmatrix}$$

দ্বিতীয় ধাপ সফলভাবে শেষ হয়েছে, তাই এবার আমাদের pivot এক ঘর সরে এসেছে নীচে ডানদিকে, মানে $(3,3)$-th entry-তে। সেখানে 0 রয়েছে। Leading 1 বানানোর জন্য তো nonzero entry চাই। তাই pivot-এর নীচে entry-গুলো দেখি। নাঃ, সেখানেও ছাই সকলেই শূন্য। তার মানে এই তৃতীয় ধাপটা ব্যর্থ হয়ে শেষ হল। এক্ষেত্রে আমরা একটা নতুন জিনিস করব, pivot-টাকে নীচে নামাব না, খালি এক ঘর ডানদিকে সরাব।

$$\longrightarrow \begin{bmatrix} 1 & 0 & 1 & 3 & -2 \\ 0 & 1 & 2 & \frac{13}{2} & -\frac{9}{2} \\ 0 & 0 & 0 & \boxed{1} & -1 \\ 0 & 0 & 0 & -\frac{3}{2} & \frac{3}{2} \end{bmatrix}$$

সৌভাগ্যক্রমে এখানে একটা nonzero entry আছে[7], এবং শুধু তাই নয়, সেটা আবার 1-ই আছে। তাই আর ভাগ করার কষ্টটা করতে হল না। খালি ওর উপরে নীচে সবাইকে 0 করলেই হবে।

$$\xrightarrow{\substack{newR_1 = R_1 - 3R_3 \\ newR_2 = R_2 - \frac{13}{2}R_3 \\ newR_4 = R_4 - \frac{3}{2}R_3}} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & \boxed{0} \end{bmatrix}$$

এই ধাপটা সফলভাবে শেষ হল, সুতরাং pivot ফের নীচে ডানদিকে একঘর সরল। ফের pivot-এর জায়গায় শূন্য আছে। কিন্তু এখানে আর nonzero আনার কোনোই পথ নেই। তাই ব্যর্থ হব। এবার pivot-কে আর ডানদিকে সরানোর জায়গা নেই, তাই আমাদের algorithm এখানেই শেষ। আমরা RREF-এ পৌঁছে গিয়েছি। ■

[7]যদি তা না থাকত, তবে ফের pivot-এর নীচে খুঁজতে হত, তাতেও কাজ না হলে ফের ব্যর্থ হয়ে pivot আরো ডানদিকে সরাতাম।

এই উদাহরণে সবরকম সম্ভাবনাই দেখানো হয়েছে বটে, কিন্তু কেবলমাত্র একটা উদাহরণ দেখে একটা algorithm শেখা সহজ না। যারা আরো উদাহরণ চাও, তারা পাশের QR code-টা scan করলেই আরো উদাহরণে পৌঁছতে পারো। সেখানে একটা button পাবে, তাতে বারবার ক্লিক করে গেলেই ধাপে ধাপে বাংলায় বোঝাতে বোঝাতে Gauss-Jordan elimination করে RREF বার করে দেখানো হবে।
https://arnabc74.github.io/linalg/gjrref.html

15.3 কী করে ব্যবহার করে?

আগেই বলেছি যে, RREF থেকে নানারকম তথ্য বার করা যায়--linearly independent কিনা, generating কিনা, ইত্যাদি। এবার একেকটা উদাহরণ দিয়ে বোঝা যাক তথ্যগুলো কী করে RREF ব্যবহার করতে হয়।

15.3.1 Linear independence পরীক্ষা করা

যদি RREF-এ একটাও zero row থাকে, তবে জানবে শুরুর তালিকাটা linearly independent ছিল না। যদি সবগুলো row-ই nonzero হয়, তবে শুরুর তালিকাটা ছিল linearly independent. একটা উদাহরণ দেখি।

Example 59: এই vector-গুলো কি linearly independent?

$$(1,2,3,4), \quad (2,3,7,5), \quad (9,5,40,-3).$$

SOLUTION: প্রথমে একটা matrix বানিয়ে নিই--

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 7 & 5 \\ 9 & 5 & 40 & -3 \end{bmatrix}.$$

এর উপর Gauss-Jordan elimination লাগালে পাবে এই RREF-টা--

$$\begin{bmatrix} 1 & 0 & 5 & -2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

যেহেতু এখানে একটা zero row এসে গেছে, তাই বলতে পারি যে গোড়ার তিনটে vector নিশ্চয়ই linearly independent ছিল না। ■

Exercise 57: এই vector-গুলো কি linearly independent?

$$(1,2,3,5), \quad (2,9,7,5), \quad (-9,0,4,13).$$

■

15.3.2 Span-এর basis আর dimension বার করা

কিছু vector-এর একটা তালিকা দেওয়া আছে, তোমার কাজ হল তাদের span-এর একটা basis বার করা। তাহলে প্রথমে যথারীতি RREF-টা বার করে নাও। কোনো zero row এসে থাকলে সেগুলোকে ফেলে দাও। খালি nonzero row-গুলো মিলে একটা মূল vector-গুলোর span-এর একটা basis হবে।

Example 60: আগের অংকের vector-গুলো নিয়েই এবারও কাজ করব। ওদের span-এর একটা basis বার করো। সেই সঙ্গে span-এর dimension-টাও কত হবে বলো।

SOLUTION: এই row echelon form-টা তো আগের অংকেই বার করেছিলাম--

$$\begin{bmatrix} 1 & 0 & 5 & -2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

এখানে nonzero row-রা হল

$$(1,0,5,-2), \quad (0,1,-1,3).$$

এরা দুজনে মিলেই গোড়ার তিনটে vector-এর span-এর basis হবে। যেহেতু basis-এ খালি 2-খানা vector আছে, তাই dimension-টা হবে 2. ■

Exercise 58: আমরা contraction theorem শেখার সময়ে শিখেছিলাম যে, যদি কিছু vector-এর একটা তালিকা দেওয়া থাকে, তবে সেই তালিকার কিছু vector নিয়েই তালিকাটার span-এর একটা basis বানানো যায়। তার জন্য "পরের পয়সায় বসে খাওয়া" vector-দের ঝেঁটিয়ে বিদায় করে দিতে হয়। এই কাজটা কী করে Gauss-Jordan elimination আর RREF দিয়ে করা যায়?

■

15.3.3 Generating কিনা পরীক্ষা করা

এবার ধরো তোমাকে কিছু vector-এর একটা তালিকা দিয়ে দেখতে বলা হল ওরা পুরো vector space-টাকে generate করে কি না। সেটা row echelon form থেকে অনায়াসে এইভাবে বার করা যায়--প্রথমে দেখো vector-গুলোর span-এর dimension কত হয় (মানে row echelon form-এ কত গুলো nonzero row আছে)। যদি সেটা পুরো vector space-টার dimension-এর সমান হয়, তবেই generate করবে, নইলে করবে না।

Example 61: চারটে vector দিলাম--

$$(1,2,3), \ (2,3,4), \ (4,5,6), \ (3,2,1).$$

এরা সবাই $\mathbb{R}^3$-এর সদস্য। এরা কি $\mathbb{R}^3$-কে generate করে?

SOLUTION: তার মানে প্রশ্নটা হল এদের span-এর dimension-টা 3 হবে কিনা। সেই প্রশ্নের উত্তর RREF বার করলেই বেরিয়ে যাবে--

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

এখানে nonzero row-এর সংখ্যা $2 < 3$. সুতরাং generating হল না। ■

Exercise 59: এই চারটে vector মিলে কি $\mathbb{R}^3$-কে span করে--

$$(1,2,7), \ (2,3,14), \ (4,15,6), \ (-3,2,1)?$$

■

15.3.4 Span-এর মধ্যে একটা vector আছে কিনা দেখা

ধরো তোমাকে কিছু vector-এর একটা তালিকা দিলাম S, আর একটা নতুন vector দিলাম $\vec{v}$. প্রশ্ন হল $\vec{v} \in \text{span}(S)$ হয় কিনা।
এর সহজ একটা কায়দা এরকম--প্রথমে একটা matrix বানাও তালিকার vector-গুলোর নীচে নতুন vector-টাকে রেখে। এবার Gauss-Jordan elimination করে RREF বার করা শুরু কর। প্রতি ধাপে এটাও হিসেব রাখো তালিকার কোন্ vector-টা কোথায় যাচ্ছে (মানে swap-গুলোর হদিস রাখো)। যখন Gauss-Jordan elimination করা শেষ হবে, তখন যদি $\vec{v}$-র row-টা একটা zero row হয়ে যায়, তবেই খালি $\vec{v} \in \text{span}(S)$ হবে, নইলে নয়।

Example 62: এই একটা vector দিলাম $(2,3,4,5)$. পরীক্ষা করে দেখো এটা এই তিনটে vector-এর span-এর মধ্যে পড়ে কিনা--

$$(1,2,3,4),\ \ (3,4,5,6),\ \ (4,5,6,7).$$

SOLUTION: তালিকার তিনটে vector-কে $\vec{v}_1$, $\vec{v}_2$ আর $\vec{v}_3$ নাম দিলাম, আর নতুন vector-টাকে বললাম $\vec{v}$. প্রথম কাজ হল এই তিনটে row হিসেবে একটার উপর একটা লিখে একটা matrix বানানো, এবং একেবারে নীচে নতুন vector-টা জুড়ে দেওয়া। এইভাবে পেলাম এই matrix-টা--

$$\begin{matrix}\vec{v}_1\\ \vec{v}_2\\ \vec{v}_3\\ \vec{v}\end{matrix}\begin{bmatrix}1&2&3&4\\3&4&5&6\\4&5&6&7\\2&3&4&5\end{bmatrix}$$

লক্ষ্য করো, row-গুলোর হদিস রাখার জন্য বাঁদিকে vector-গুলোর নাম লেবেলের মত লিখে নিয়েছি। এবার Gauss-Jordan elimination দিয়ে RREF বার করলে পাব--

$$\begin{matrix}\vec{v}_1\\ \vec{v}_2\\ \vec{v}_3\\ \vec{v}\end{matrix}\begin{bmatrix}1&0&-1&-2\\0&1&2&3\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

এখানে কোনো swap লাগেনি, তাই বাঁদিকের লেবেলগুলো একই রয়েছে। লক্ষ্য করো, $\vec{v}$-এর row-টা এখন একটা zero row. তাই সিদ্ধান্ত করতে পারি যে নতুন vector-টা বাকিদের span-এর মধ্যে ছিল। ■

এবার আরেকটা উদাহরণ দেখি যেখানে নতুন vector-টা span-এর বাইরে পড়বে।

Example 63: এখানেও আগের তালিকাটাই নেব। খালি নতুন vector-টা নেব $(2,3,5,5)$. এবার দ্যাখো span-এর মধ্যে পড়ে কিনা।

SOLUTION: এখানে Gauss-Jordan elimination করে RREF-টা হচ্ছে এইরকম--

$$\begin{matrix}\vec{v}_1\\ \vec{v}_2\\ \vec{v}\\ \vec{v}_3\end{matrix}\begin{bmatrix}1&0&0&-2\\0&1&0&3\\0&0&1&0\\0&0&0&0\end{bmatrix}$$

এখানে $\vec{v}$-এর row-টা একটা nonzero row. সুতরাং $\vec{v}$-টা $\text{span}\{\vec{v}_1,\vec{v}_2,\vec{v}_3\}$-এর মধ্যে রইল না। ■

Exercise 60: পরীক্ষা করে দেখো $(1,2,3,4) \in \text{span}\{(3,4,5,6),(-1,0,4,5),(4,5,2,1)\}$ কিনা। ■

15.3.5 Basis extension

ধরো একটা linearly independent তালিকা দিয়ে সেটাকে পুরো vector space-টার একটা basis-এ extend করতে বলল। তখন তালিকাটার RREF বার করে দেখতে হবে কোন্ কোন্ column-এ কোনো leading 1 নেই। ধরো দেখলে প্রথম, তৃতীয় আর চতুর্থ column-এ কোনো leading 1 নেই। তবে তালিকার সঙ্গে $\vec{e}_1, \vec{e}_3$ আর $\vec{e}_4$ জুড়ে দিলেই একটা basis তৈরী হবে।

Example 64: ধরো তালিকাটা নিলাম $(1,2,3,4)$, $(2,4,10,9)$. এটাকে $\mathbb{R}^4$-এর একটা basis-এ extend করো।

SOLUTION: এখানে RREF-টা হবে

$$\begin{bmatrix} \mathbf{1} & 2 & 0 & 1 \\ 0 & 0 & \mathbf{1} & 1 \end{bmatrix}$$

এখানে leading 1-গুলোকে মোটা করে দেখিয়েছি। লক্ষ্য করো, leading 1 নেই দ্বিতীয় আর চতুর্থ column-এ। তাই extend করার জন্য $\vec{e}_2$ আর $\vec{e}_4$ লাগবে, মানে extend করে আমরা এই basis-টা পাব--

$$\{(1,2,3,4),(2,4,10,9),(0,1,0,0),(0,0,0,1)\}.$$

■

Exercise 61: Extend $\{(1,4,8,2),\ (3,4,-7,0)\}$ to a basis of $\mathbb{R}^4$. ■

Answers

1. $\mathbb{R}$ নিজেই, তার মানে $\mathbb{R}$ নিজেই নিজের উপরে একটা vector space. এখানে real number-রা একাধারে vector এবং scalar-এর ভূমিকা পালন করছে। **2.** দ্বিতীয়টা, কারণ একটা complex number-কে real number দিয়ে গুণ করলে উত্তরটা complex-ই থাকে। কিন্তু একটা real number-কে complex number দিয়ে গুণ করলে উত্তরটা real নাও থাকতে পারে। তাই প্রথম ক্ষেত্রে scalar multiplication-এর সংজ্ঞা নিয়ে সমস্যা হবে। **4.** $0 \cdot \vec{v}$ মানে যে $\vec{0}$ সেটা প্রমাণ করতেই তো additive inverse-এর অস্তিত্ব ব্যবহার করতে হয়। **7.** (A). যেহেতু $(0,0,0)$ এখানে (B) আর (C)-এর শর্ত পালন করছে না, তাই ওরা বাদ। (A)-র বেলায় closed under addition and scalar multiplication হচ্ছে। তাই (A) হবে, এবং (D) হবে না। **10.** যেহেতু $S \neq \phi$, তাই কোনো একটা $\vec{v} \in S$ আছে। সুতরাং $0\vec{v}$-ও S-এর মধ্যেই আছে, কারণ closed under scalar multiplication বলে দিয়েছে। এদিকে আমরা জানি যে $0\vec{v} = \vec{0}$. সুতরাং $\vec{0} \in S$ প্রমাণ হয়ে গেল। **12.** Field-টা নাও $\mathbb{R}$, আর V-ও নাও $\mathbb{R}$. আর S নিতে পারো $[0, \infty)$. চাইলে $S = \mathbb{Z}$-ও নিতে পারো। **13.** $V = \mathbb{R}^2$ আর $S = \{(x,y)\ :\ x = 0 \text{ or } y = 0\}$ নাও। **14.** যেহেতু $\vec{u}, \vec{v} \in \text{span}(S)$, তাই ওদেরকে নিশ্চয়ই এভাবে linear combination হিসেবে লেখা যাবে--

$$\begin{aligned} \vec{u} &= p_1\vec{u}_1 + \cdots + p_m\vec{u}_m \\ \vec{u} &= q_1\vec{v}_1 + \cdots + q_n\vec{v}_n \end{aligned}$$

যেখানে $p_i, q_j \in F$ এবং $\vec{u}_i, \vec{v}_j \in S$. সুতরাং

$$a\vec{u} + b\vec{v} = \sum_1^m ap_i\vec{u}_i + \sum_1^n bq_i\vec{v}_j,$$

যেটা আবার S-এর vector-দেরই একটা linear combination হচ্ছে। সুতরাং $a\vec{u} + b\vec{v} \in \text{span}(S)$. **15.** দেখতে হবে $(5,2,1) = a(1,2,3) + b(3,1,0)$ লেখা যায় কিনা। এরকম a, b পাওয়ার জন্য এই তিনটে equation-কে solve করতে হবে, $a + 3b = 5,\ 2a + b = 2$ এবং $3a = 1$. শেষেরটা থেকে $a = \frac{1}{3}$. প্রথমটায় বসিয়ে $b = \frac{14}{9}$. দ্বিতীয়টায় বসালে মিলছে না। তার মানে এরকম a, b সম্ভব নয়। সুতরাং $(5,2,1) \notin \text{span}\{(1,2,3),(3,1,0)\}$. **16.** span$(S)$-এর সব সদস্যই তো S-এর বিভিন্ন linear combination. যেহেতু $S \subseteq U$, তাই S-এর linear combination-রাও U-এর ভিতরেই থাকবে। **17.**

> Solving $(1,1,0,0) = a(2,-1,3,3) + b(0,1,-1,-1)$, we get $a = \frac{1}{2}$ and $b = \frac{3}{2}$.
> Again, solving $(1,0,1,1) = c(2,-1,3,3) + d(0,1,-1,-1)$, we get $c = d = \frac{1}{2}$.
> So $\{(1,1,0,0),\ (1,0,1,1)\} \subseteq \text{span}\{(2,-1,3,3),\ (0,1,-1,-1)\}$.
> Hence $\text{span}\{(1,1,0,0),\ (1,0,1,1)\} \subseteq \text{span}\{(2,-1,3,3),\ (0,1,-1,-1)\}$.

এইটা কেন হল? কারণ $\text{span}\{(2,-1,3,3),\ (0,1,-1,-1)\}$ হল একটা subspace. তার মধ্যে যেহেতু $(1,1,0,0)$ আর $(1,0,1,1)$ রয়েছে, তাই ওদের যাবতীয় linear combination-ও থাকবে।

> Again, solving $(2,-1,3,3) = a(1,1,0,0) + b(1,0,1,1)$, we get $a = -1$ and $b = 3$.
> Also, solving $(0,1,-1,-1) = c(1,1,0,0) + d(1,0,1,1)$, we get $c = 1$ and $d = -1$.
> So $\{(2,-1,3,3),\ (0,1,-1,-1)\} \subseteq \text{span}\{(1,1,0,0),\ (1,0,1,1)\}$.
> Hence $\text{span}\{(2,-1,3,3),\ (0,1,-1,-1)\} \subseteq \text{span}\{(1,1,0,0),\ (1,0,1,1)\}$.
> Thus, $\text{span}\{(2,-1,3,3),\ (0,1,-1,-1)\} = \text{span}\{(1,1,0,0),\ (1,0,1,1)\}$, as required.

18. $\sum_i(\alpha_i - \beta_i)\vec{v}_i = \vec{0}$. যেহেতু linearly independent, তাই প্রতিটা coefficient-ই আলাদা করে শূন্য হতে বাধ্য, মানে $\forall i \quad \alpha_i - \beta_i = 0$. **19.** ধরো $\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = \vec{0}$. আমরা দেখাব যে $\alpha_1 = \cdots = \alpha_n = 0$. এদিকে যদি $\beta_1 = \cdots = \beta_n = 0$ নাও, তবে $\beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n = \vec{0}$ হবে। সুতরাং $\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = \beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n$ হচ্ছে। অতএব $\forall i \quad \alpha_i = \beta_i = 0$ হবে। **20.** ধরো $a_1\vec{v}_1 + \cdots + a_n\vec{v}_3 = \vec{0}$. যেহেতু $\vec{v}_n \notin \text{span}\{\vec{v}_1, ..., \vec{v}_{n-1}\}$, তাই $a_n = 0$. সুতরাং $a_1\vec{v}_1 + \cdot + a_{n-1}\vec{v}_{n-1} = \vec{0}$. যেহেতু $\vec{v}_{n-1} \notin \text{span}\{\vec{v}_1, ..., \vec{v}_{n-2}\}$, তাই $a_{n-1} = 0$. পরের ধাপটা আশা করি বুঝতেই পারছ। গুছিয়ে লিখতে induction লাগবে। **21.** না, যেমন $\mathbb{R}^3$-র মধ্যে $A = \{(1,0,0),(0,1,0)\}$ আর $B = \{(1,1,1),(1,1,-1)\}$ নাও। তাহলে A, B প্রশ্নের সব শর্তই পালন করে, কিন্তু তাও $A \cup B$ মোটেই linearly independent নয়, কারণ $(1,1,1) = 2(1,0,0) + 2(0,1,0) - (1,1,-1)$. **22.** $\dim(S) = 2$. **23.** $\dim(W) = 2$. একটা basis হল $\{(1,-1,0),(0,1,-1)\}$. **24.** একটা basis হল $\{(1,-1,-1)\}$ এবং $\dim(S) = 1$. **25.** $\dim(W) = 2$. **26.** একটা basis হল $\left\{\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$. **27.** একটা basis হল $\left\{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$. এখানে dimension হল 3. **28.** **(1)** $(-3,-1)$, **(2)** $(-3,2)$, **(3)** $(-1,3)$, **(4)** $(3/2,4)$, **(5)** $(4,-3)$, **(6)** $(-4,-3)$, **(7)** $(-3,-4)$.
29. $(5,6,16)$. **30.** এখানে এই equation-গুলোকে solve করতে হবে, $a+b+c=1,\ a+b=3$ এবং $a=1$. উত্তর হল $(a,b,c) = (1,2,-2)$. **31.** 27. **32.** **(1)** 2, কারণ খালি x_1, x_4 জানলেই চলবে। **(2)** 3, কারণ $Re(z_1), Im(z_1), Re(z_2)$ জানলেই হবে। **33.** এটা $\mathbb{C}$-এর উপরে একটা vector space-ই নয়! যেমন ধরো $\vec{v} = (1+i, i) \in V_2$ কিন্তু যদি $\alpha = i$ নিই তবে $\alpha\vec{v} = (-1+i, -1) \notin V_2$. **34.** (B). এর সংক্ষিপ্ত প্রমাণ হল--

A matrix is in W if and only if it is of the form

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} = a\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + e\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + f\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

So these 6 matrices span W. Also these are linearly independent. So W has a basis of size 6.

35. (B). এর সংক্ষিপ্ত justification হল এইরকম--

A matrix is in U if and only if it is of the form

$$\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} = c\underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{A} + d\underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}}_{B}.$$

So $\{A, B\}$ spans U. Also $\{A, B\}$ is linearly independent. So $\{A, B\}$ is a basis for U.

Similarly, a matrix is in W if and only if it is of the form

$$\begin{bmatrix} a & b \\ c & -c \end{bmatrix} = a\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{P} + b\underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{Q} + c\underbrace{\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}}_{R}.$$

So $\{P, Q, R\}$ spans V. Also $\{P, Q, R\}$ is linearly independent. So $\{P, Q, R\}$ is a basis for V.

36. এখানে $(1,3,2) = (1,1,0) + 2(0,1,1)$. তাই $(1,3,2)$ দিয়ে $(1,0,1)$-কে replace করা যাবে না, কারণ তাহলে পাব $\{(1,1,0),\ (0,1,1),\ (1,3,2)\}$, যেটা linearly dependent. **37.** True. **38.** $\mathbb{R}^3$-র dimension হল 3. মোটে দুটো vector দিয়ে কখনোই ওর basis তৈরী হতে পারে না! **39.** একটা সম্ভাব্য উত্তর হল $\{(1,1,-1),\ (2,1,1),\ (1,0,0)\}$. **40.** ভুল। একটা counterexample: $V = \mathbb{R}^2, n = 1, \vec{u}_1 = (1,1), \vec{v}_1 = (1,0), \vec{v}_2 = (0,1)$. তাহলে $(1,0),(0,1) \not\in \text{span}\{(1,1)\}$ কিন্তু তা বলে $\{(0,1),(1,0),(1,1)\}$ অবশ্যই linearly independent নয়, কারণ $(1,1) = (1,0) + (0,1)$. **41.**

Every generating set must have size $\geq \dim(\mathbb{R}^3) = 3$.

42. $(-2,1,1) \in U \cap W$. আসলে $W = \text{span}\{(-2,1,1)\}$, যেটা একটা লাইন। **43.** হ্যাঁ, কারণ union-টা xy-plane নিজেই। **44.** ধরে নিয়েছি $U \not\subseteq W$. তার মানে U-এর মধ্যে এমন কিছু আছে, যেটা W-র মধ্যে নেই। এরকম একটা vector নিয়ে শুরু করব--

> Only if: Let $\vec{u} \in U \setminus W$.
>
> Take any $\vec{w} \in W$. Shall show $\vec{w} \in U$.

এতেই দেখানো হয়ে যাবে যে, $W \subseteq U$.

> $\vec{u}, \vec{w} \in U \cup W$. Since $U \cup W$ is a subspace, hence $\vec{u} + \vec{w} \in U \cup W$.
>
> Case I: $\vec{u} + \vec{w} \in U$: Then $\vec{w} = (\vec{u} + \vec{w}) - \vec{u}) \in U$, as required.
>
> Case II: $\vec{u} + \vec{w} \notin U$: Then $\vec{u} + \vec{w} \in W$. So $\vec{u} = (\vec{u} + \vec{w}) - \vec{w}) \in U$, which is not possible by the given condition. So this case is impossible.

45. পুরো $\mathbb{R}^3$. **46.** **(1)** yz-plane এবং $\{(x, y, z) \; : \; x = z\}$. **(2)** z-axis এবং $\text{span}\{(1, 1, 1)\}$. **(3)** xy-plane এবং xz-plane. **(4)** পুরো $\mathbb{R}^3$-ই হল একমাত্র direct complement. **(5)** $\{\vec{0}\}$ হল একমাত্র direct complement.
47. দুটোই ভুল। প্রথমটা ভুল কারণ $(0, 0, 1) \notin \{(x, 0, 0) \; : \; x \in \mathbb{R}\} + \{(0, y, 0) \; : \; y \in \mathbb{R}\}$. দ্বিতীয়টা ভুল কারণ ওদের intersection-টা $\{\vec{0}\}$ নয়, আস্ত x-axis-টা। **48.**

> If $\vec{v} = \vec{s}_1 + \vec{t}_1 = \vec{s}_2 + \vec{t}_2$, then $\vec{s}_1 - \vec{s}_2 = \vec{t}_2 - \vec{t}_1$, which is both S and T. So Since $S \cap T = \{\vec{0}\}$, hence $\vec{s}_1 = \vec{s}_2$ and $\vec{t}_1 = \vec{t}_2$.

49.

> Since every $\vec{v} \in V$ can be written as $\vec{v} = \vec{s} + \vec{t}$ with $\vec{s} \in S$ and $\vec{t} \in T$, we have $V = S + T$.
>
> Let, if possible, $S \cap T \neq \{\vec{0}\}$.
>
> We pick any nonzero $\vec{w} \in S \cap T$. Then any $\vec{v} \in V$ can be split up in at least two distinct ways: $\vec{v} = \vec{s} + \vec{t}$ and $\vec{v} = (\vec{s} + \vec{w}) + (\vec{t} - \vec{w})$ contradicting the given condition.

50. S এবং $(1, 1, 0) + S$ একই জিনিস। **51.** $\mathbb{R}^7$-এর মধ্যে $S = \text{span}\{\vec{e}_1, ..., \vec{e}_5, \vec{e}_6\}$ এবং $T = \text{span}\{\vec{e}_1, ..., \vec{e}_5, \vec{e}_7\}$. তবে $S \cap T = \text{span}\{\vec{e}_1, ..., \vec{e}_5\}$. **52.** $\dim(U) = \dim(W) = 3$, $\dim(U \cap W) = 2$. সুতরাং $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W) = 3 + 3 - 2 = 4$.
53. $7 + 9 - 10 = 6$. এর চেয়ে নীচে $\dim(S \cap T)$ নামতে পারে না। কারণ যদি $\dim(S \cap T) < 6$ হয়, তবে $\dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T) > 10$ হয়ে যাচ্ছে, যেটা অসম্ভব, কারণ $S + T$ হল $\mathbb{R}^{10}$-এর একটা subspace. আবার এমন S, T সম্ভব, যাতে $\dim(S \cap T) = 6$ হয়। যেমন $S = \text{span}\{\vec{e}_1, ..., \vec{e}_6, \vec{e}_7\}$ আর $T = \text{span}\{\vec{e}_1, ..., \vec{e}_6, \vec{e}_8, \vec{e}_9, \vec{e}_{10}\}$. **54.** 1. **55.** 2. **56.** $n - 2$. আগের অংকের যুক্তিতে dimension-টা এর নীচে নামতে পারে না। এর উপরেও উঠতে পারে না, কারণ সেক্ষেত্রে $U = V$ হয়ে যাচ্ছে, যেহেতু $\dim(U) = \dim(V) = n - 1$. **57.** হ্যাঁ, কারণ RREF-টা হল $\begin{bmatrix} 1 & 0 & 0 & \frac{-29}{137} \\ 0 & 1 & 0 & \frac{-213}{137} \\ 0 & 0 & 1 & \frac{380}{137} \end{bmatrix}$. **58.** Gauss-Jordan elimination দিয়ে যখন RREF বার করা হয়, তখন প্রতিটা row-ই আসে শুরু তালিকার কোনো না কোনো vector থেকে। Contraction theorem লাগানোর জন্য কোন্ row-টা কোন্ vector থেকে এসেছে সেটা খেয়াল রাখতে হয় (বিশেষ করে swap ধাপগুলোতে)। শেষপর্যন্ত RREF-এর nonzero row-গুলো শুরুর তালিকার যে যে vector থেকে এসেছে, সেগুলো মিলেই একটা basis হবে। **59.** Vector চারটেকে row হিসেবে নিয়ে যে 4×3 সাইজের matrix-টা

হয়, তার RREF হচ্ছে $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. সুতরাং এদের span-এর dimension হল 3. তার মানে ওরা মিলে $\mathbb{R}^3$-কে span করে। **60.** এই চারটে vector দিয়ে তৈরী matrix-টার RREF হচ্ছে I. তার মানে vector-গুলো কেউ কারো span-এ নেই। **61.** এখানে RREF-টা আসছে $\begin{bmatrix} 1 & 0 & -\frac{15}{2} & -1 \\ 0 & 1 & \frac{31}{8} & \frac{3}{4} \end{bmatrix}$. যেহেতু তৃতীয় আর চতুর্থ column-এ কোনো leading 1 নেই, তাই $\vec{e}_3, \vec{e}_4$ জুড়ে দিলেই basis-এ extend করা হবে।

Chapter III
Matrix

DAY 16
$\mathscr{C}(A), \mathscr{R}(A)$ আর $\mathscr{N}(A)$ (part 1)

আমরা যখন প্রথম matrix-এর কথা শিখি তখন ওদেরকে মনে করি একগুচ্ছ সংখ্যা হিসেবে, যারা একটা rectangle-এর মত করে সাজানো। কিন্তু এভাবে চিন্তা করে বেশীদূর এগোনো মুস্কিল। তার চেয়ে ভালো কায়দা হল একটা matrix-কে এক গুচ্ছ row (বা একগুচ্ছ column) বলে ভাবা। এখানে প্রত্যেকটা row (বা column) হল একেকটা vector.
ধরো একটা $m \times n$ matrix দেওয়া আছে--

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 9 & -4 \end{bmatrix}.$$

এখানে $m = 2$ আর $n = 3$. এই matrix-টাকে আমরা চাইলে $n = 3$-টে column দিয়ে তৈরী বলে ভাবতে পারি--

$$A = [\ \vec{a}_1 \quad \vec{a}_2 \quad \vec{a}_3\ ],$$

যেখানে

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 0 \\ 9 \end{bmatrix} \text{ আর } \vec{a}_3 = \begin{bmatrix} 2 \\ -4 \end{bmatrix}.$$

আবার চাইলে তুমি একই matrix-কে $m = 2$ খানা row দিয়ে তৈরী বলেও ভাবতে পারো--

$$A = \begin{bmatrix} \vec{b}_1' \\ \vec{b}_2' \end{bmatrix},$$

যেখানে

$$\vec{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \text{ আর } \vec{b}_2 = \begin{bmatrix} 3 \\ 9 \\ -4 \end{bmatrix}.$$

লক্ষ্য করো, row-রা যদিও "শোয়ানো" vector, কিন্তু আলাদা করে লেখার সময়ে যেকোনো vector-এর মত ওদেরও খাড়া করে লিখেছি। কিন্তু ভালো করে দ্যাখো A-কে লেখার সময়ে কিন্তু $\vec{b}_1'$ আর $\vec{b}_2'$ লিখেছি, মানে $\vec{b}_1, \vec{b}_2$-এর transpose. ওই transpose করার ফলেই বোঝা যাচ্ছে যে A-র মধ্যে ওরা "শুয়ে" আছে।
লক্ষ্য করো, $\vec{b}_i$-রা সবাই $\mathbb{R}^n$-এর সদস্য, আর $\vec{a}_j$-রা $\mathbb{R}^m$-এর। সুতরাং $\mathrm{span}\{\vec{a}_1, ..., \vec{a}_n\}$ হবে $\mathbb{R}^m$-এর একটা subspace, এবং $\mathrm{span}\{\vec{b}_1, ..., \vec{b}_m\}$ হবে $\mathbb{R}^n$-এর subspace. প্রথম subspace-টা column-গুলো দিয়ে generated, তাই ওকে বলব A-র **column space**, আর লিখব $\mathscr{C}(A)$. একই কারণে অন্য subspace-টার নাম দেব A-র **row space**, আর লিখব $\mathscr{R}(A)$.

এবার সেই সব vector $\vec{x}$ নাও, যাদের জন্য $A\vec{x} = \vec{0}$ হয়। $A\vec{x}$ বার করার জন্য অবশ্যই $\vec{x} \in \mathbb{R}^n$ হতে হবে। একটু চিন্তা করলেই দেখবে যে

$$\{\vec{x} \in \mathbb{R}^n \ : \ A\vec{x} = \vec{0}\}$$

হল $\mathbb{R}^n$-এর একটা subspace. কারণ--

> যদি $\vec{x}, \vec{y}$ এমন হয়, যাতে $A\vec{x} = A\vec{y} = \vec{0}$, আর a, b যে কোনো দুটো scalar হয়, তবে
>
> $$A(a\vec{x} + b\vec{y}) = aA\vec{x} + bA\vec{y} = a\vec{0} + b\vec{0} = \vec{0}.$$
>
> তার মানে $a\vec{x} + b\vec{y}$-এরও সেই একই বৈশিষ্ট্য থাকছে, A দিয়ে বাঁদিক থেকে গুণ করলে $\vec{0}$ হয়ে যায়।

এই subspace-টাকে বলে A-র **nullspace** বা **kernel.** লেখার সময়ে লিখি $\mathscr{N}(A)$.
এই যে কয়েকটা subspace-এর কথা শিখলাম, এদের মধ্যে কয়েকটা সম্পর্ক আছে যেগুলো অত্যন্ত গুরুত্বপূর্ণ। এবার এদের সম্পর্কে আলোচনা করব। প্রমাণের কথা পরে ভাবা যাবে, আগে জিনিসগুলো বুঝে নাও।
নীচের অংকে একটা matrix দিয়ে তার row space বার করতে দিয়েছে। Row space মানে তো একটা vector space, সেটা "বার করা" বলতে কী বোঝায়? একটা উত্তর হল vector space-টার একটা basis বার করা। আমরা সেভাবেই করব অংকটা।

Example 1: Find the row space of

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 3 & 0 & 7 \\ -1 & 4 & 3 \end{bmatrix}.$$

Examine if $(1,1,1)$, $(1,-1,1)$ are in the row space of A.[3] **(2012.4bi)**
SOLUTION: এখানে row-গুলো হল একগুচ্ছ vector, তাদের span-এর একটা basis বার করতে হবে। এই রকম অংক আমরা দ্বিতীয় অধ্যায়েই করেছি, RREF দিয়ে। অবশ্য এরকম ছোটো অংক RREF ছাড়াই হয়ে যায়--

> Here the first two rows are not multiples of each other.

ব্যস্ ওরা অন্ততঃ linearly independent. এবার রাফে খানিকক্ষণ চেষ্টা করে দেখে নাও তৃতীয় row-টাকে ওদের linear combination হিসেবে লেখা যাচ্ছে কিনা, মানে এমন $a, b \in \mathbb{R}$ আছে কিনা যাতে $a(1,2,5) + b(3,0,7) = (-1,4,3)$ হয়। তার মানে $a + 3b = -1$, $2a = 4$ এবং $5a + 7b = 3$. দ্বিতীয়টা থেকে $a = 2$ এবং প্রথমটা থেকে $b = -1$ হবে। এবার দেখতে হবে তৃতীয় equation-টাতে $a = 2, b = -1$ বসালে মেলে কিনা। যদি মেলে, linear combination হিসেবে লেখা যাবে, নইলে যাবে না। এক্ষেত্রে মিলে যাচ্ছে, কারণ $5 \times 2 + 7(-1) = 3$. সুতরাং--

> The third row lies in the span of the first two, since
>
> $$2(1,2,5) - (3,0,7) = (-1,4,3).$$
>
> So $\mathscr{R}(A)$ is the space with basis $\{(1,2,5),(3,0,7)\}$.

এবার দেখতে হবে $(1,1,1)$ এবং $(1,-1,1)$ এর মধ্যে পড়ে কি না। ফের equation বানানো যাক।

> Next, we check if $(1,1,1) = c(1,2,5) + d(3,0,7)$ for some $c, d \in \mathbb{R}$,
> ie, $c + 3d = 1,\ 2c = 1$ and $5c + 7d = 1$.
> From the second equation $c = \frac{1}{2}$ and from the first $d = \frac{1}{6}$.

> But these do not satisfy the third equation. Hence $(1,1,1) \notin \mathscr{R}(A)$.

একইভাবে অন্য vector-টার বেলাতেও--

> Next, we check if $(1,-1,1) = p(1,2,5) + q(3,0,7)$ for some $p, q \in \mathbb{R}$,
> ie, $p + 3q = 1,\ 2p = -1$ and $5p + 7q = 1$.
> From the second equation $p = -\frac{1}{2}$ and from the first $q = \frac{1}{2}$.
> These also satisfy the the third equation. Hence $(1,-1,1) \in \mathscr{R}(A)$.

যদি RREF ব্যবহার করতে তবে সবকিছু একবারে হয়ে যেত। সেটা সংক্ষেপে করে দেখাই।

বিকল্প পদ্ধতি

প্রথমে ওই দুটো vector-কে matrix-টার নীচে জুড়ে একটা বড় matrix বানাই।

$$\begin{matrix} \vec{v}_1 \\ \vec{v}_2 \\ \vec{v}_3 \\ \vec{u} \\ \vec{v} \end{matrix} \begin{bmatrix} 1 & 2 & 5 \\ 3 & 0 & 7 \\ -1 & 4 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

পাশে লেবেল লাগিয়ে দিয়েছি, $\vec{v}_1, \vec{v}_2, \vec{v}_3$ হল মূল matrix-টার row-গুলো, আর নতুন vector দুটোকে $\vec{u}$ আর $\vec{v}$ নাম দিয়েছি। এবার Gauss-Jordan elimination লাগিয়ে RREF বানাব। মাঝের ধাপগুলো আর দেখাচ্ছি না। শেষ পর্যন্ত পাব--

$$\begin{matrix} \vec{v}_1 \\ \vec{v}_2 \\ \vec{u} \\ \vec{v}_3 \\ \vec{v} \end{matrix} \begin{bmatrix} 1 & 0 & \frac{7}{3} \\ 0 & 1 & \frac{4}{3} \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

বিভিন্ন swap-এর ফলে লেবেলগুলো কেমন বদলে গেছে দেখো। এবার $\vec{v}_1, \vec{v}_2, \vec{v}_3$-এর মধ্যে যেগুলো nonzero আছে, তাদের দিয়ে $\mathscr{R}(A)$-র basis পাবে, মানে $\left\{\left(1,0,\frac{7}{3}\right), \left(0,1,\frac{4}{3}\right)\right\}$. দেখা যাচ্ছে $\vec{u}$-এর row-টা nonzero আছে, তাই $(1,1,1) \notin \mathscr{R}(A)$. কিন্তু $\vec{v}$-এর row-টা পুরো zero হয়ে গেছে, তাই $(1,-1,1) \in \mathscr{R}(A)$. ■

16.1 $\mathscr{C}(A)$ আর $\mathscr{R}(A)$-এর সদস্যদের দেখতে কেমন?

$\mathscr{C}(A)$ আর $\mathscr{R}(A)$-এর সদস্যদের একটা বিশেষ চেহারা আছে, যেটা বারবার কাজে লাগে। সেটা শিখে রাখি। ধরো এই matrix-টা নিলাম--

$$A = \begin{bmatrix} 1 & 10 \\ 2 & 100 \\ 3 & 1000 \end{bmatrix}.$$

এর দুটো column দুটোর যা খুশী একটা linear combination নাও, যেমন--

$$\vec{v} = 2\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 10 \\ 100 \\ 1000 \end{bmatrix}.$$

বলাই বাহুল্য যে $\vec{v} \in \mathscr{C}(A)$ হবে। এবার লক্ষ্য কর, $\vec{v}$-কে আমরা এইভাবে লিখতে পারি--

$$\vec{v} = \begin{bmatrix} 1 & 10 \\ 2 & 100 \\ 3 & 1000 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

একেবারে ডানদিকের $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$-টা কী ভাবে linear combination-টার coefficient-গুলো পরপর বসিয়ে তৈরী হয়েছে দেখো। এ থেকে নিশ্চয়ই বুঝতে পারছ যে, $\mathscr{C}(A)$-র যেকোনো সদস্যকেই $A\vec{x}$ আকারে লেখা যায়। এর উল্টোটাও সত্যিই--যা খুশি $\vec{x}$ নাও, তা হলে $A\vec{x}$ অবশ্যই $\mathscr{C}(A)$-র মধ্যে থাকতে বাধ্য। একটা উদাহরণ নিই।

Example 2: মনে করো A নিলাম আগেরটাই, আর $\vec{x} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$. তাহলে

$$A\vec{x} = \begin{bmatrix} 1 & 10 \\ 2 & 100 \\ 3 & 1000 \end{bmatrix} \begin{bmatrix} -2 \\ 5 \end{bmatrix} = -2 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 5 \times \begin{bmatrix} 10 \\ 100 \\ 1000 \end{bmatrix} \in \mathscr{C}(A).$$

■

এই কথাটা এতই গুরুত্বপূর্ণ যে একটা theorem-এর আকারে লিখে রাখা যাক--

> THEOREM
>
> Let $A_{m\times n}$ be any matrix. Let $\vec{v}_{m\times 1}$ be any vector. Then $\vec{v} \in \mathscr{C}(A)$ iff $\vec{v}$ can be written as $\vec{v} = A\vec{x}$ for some $\vec{x}_{n\times 1}$.

ঠিক একই রকম কথা বলা যায় $\mathscr{R}(A)$-এর সম্বন্ধেও। খালি মনে রেখো যে $\mathscr{R}(A) = \mathscr{C}(A')$ হয় (কারণ transpose নিলেই row-রা column, এবং column-রা row হয়ে যায়)। সুতরাং--

> THEOREM
>
> Let $A_{m\times n}$ be any matrix. Let $\vec{v}_{n\times 1}$ be any vector. Then $\vec{v} \in \mathscr{R}(A)$ iff $\vec{v}' = \vec{x}'A$ for some $\vec{x}_{m\times 1}$.

আমরা আগের দুটো অংকে $\mathscr{C}(A)$-র দুজন সদস্যের কথা পেয়েছি। একজন হল

$$\begin{bmatrix} 32 \\ 304 \\ 3006 \end{bmatrix} = \begin{bmatrix} 1 & 10 \\ 2 & 100 \\ 3 & 1000 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix},$$

আর অন্যজন হল

$$\begin{bmatrix} 48 \\ 496 \\ 4994 \end{bmatrix} = \begin{bmatrix} 1 & 10 \\ 2 & 100 \\ 3 & 1000 \end{bmatrix} \begin{bmatrix} -2 \\ 5 \end{bmatrix}.$$

এদের যদি পাশাপাশি বসিয়ে একটা matrix বানাই, তবে তাকে এইভাবে লেখা যাবে--

$$\underbrace{\begin{bmatrix} 32 & 48 \\ 304 & 496 \\ 3006 & 4994 \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} 1 & 10 \\ 2 & 100 \\ 3 & 1000 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 2 & -2 \\ 3 & 5 \end{bmatrix}}_{X}.$$

যেহেতু B-এর column-গুলো সবাই A-র column-গুলোর linear combination, সুতরাং এতে নিশ্চয়ই সন্দেহ নেই যে $\mathscr{C}(B) \subseteq \mathscr{C}(A)$.
অপর পক্ষে যদি $B = AX$ হয়, তবে অবশ্যই $\mathscr{C}(B) \subseteq \mathscr{C}(A)$ হতে বাধ্য। সুতরাং আরেকটা theorem পাওয়া গেল--

THEOREM

Let $A_{m\times n}$ and $B_{m\times p}$ be two matrices. Then $\mathscr{C}(B) \subseteq \mathscr{C}(A)$ iff $B = AX$ for some $X_{n\times p}$.

বলাই বাহুল্য যে row space-এর জন্যও এর একটা সংস্করণ আছে--

THEOREM

Let $A_{m\times n}$ and $B_{p\times n}$ be two matrices. Then $\mathscr{R}(B) \subseteq \mathscr{R}(A)$ iff $B = XA$ for some $X_{p\times m}$.

এই দুটো theorem-কে এভাবে মনে রাখতে পারো--$\mathscr{C}(B) \subseteq \mathscr{C}(A)$ হওয়া মানে A-কে ডানদিক থেকে কিছু একটা দিয়ে গুণ করলে B পাওয়া যায়। আর $\mathscr{R}(B) \subseteq \mathscr{R}(A)$ মানে গুণটা বাঁদিক থেকে করতে হবে।

16.2 $\mathscr{C}(A)$ আর $\mathscr{R}(A)$-এর সম্পর্ক

এবার $\mathscr{C}(A)$ আর $\mathscr{R}(A)$-র সম্পর্কের কথা বলি।
একটা সম্পর্ক তো বোঝাই যায়--যদি একটা matrix-এর transpose নাও, তবে row-গুলো হয়ে যায় column, আর column-গুলো হয়ে যায় row. সুতরাং যেকোনো matrix A-র জন্যই

$$\mathscr{C}(A) = \mathscr{R}(A') \text{ এবং } \mathscr{R}(A) = \mathscr{C}(A').$$

এইটার মধ্যে আশ্চর্য হবার কিছুই নেই। কিন্তু যেটা রীতিমত আশ্চর্য ব্যাপার, সেটা হল--

$$\dim(\mathscr{C}(A)) = \dim(\mathscr{R}(A)).$$

$\mathscr{C}(A)$ আর $\mathscr{R}(A)$ সম্পূর্ণ আলাদা দুটো subspace. যদি A হয় $m \times n$, তবে $\mathscr{C}(A)$ রয়েছে $\mathbb{R}^m$-এর মধ্যে, আর $\mathscr{R}(A)$ রয়েছে $\mathbb{R}^n$-এর মধ্যে, অথচ কি আজব কাণ্ড, দুজনের dimension সমান হয়ে বসে আছে! অংকের দুনিয়ায় এরকম আজব ঘটনা মাঝে মাঝে হয়, এবং তার ফলে জীবন অনেক সহজ হয়ে যায়। একটা কথা মনে পড়ে গেল, একটু অপ্রাসঙ্গিক হলেও বলি। বিখ্যাত বৈজ্ঞানিক Feynman নাকি বলতেন যে, অংকের দুনিয়াতে অন্যতম আশ্চর্য জিনিস হল

$$e^{i\pi} = -1.$$

সত্যিই তো, কোথায় ছিল $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots$, আর কোথায় ছিল π যেটা কিনা একটা circle-এর circumference (পরিধি)-র সঙ্গে তার diameter (ব্যাস)-এর ratio, আর কোথায় ছিল complex সংখ্যা i, এই তিন জগতের তিনজন কী করে বেমালুম মিলে -1 হয়ে গেল?

আবার linear algebra-র প্রসঙ্গে ফিরে আসি। এক্ষুণি যেটা বললাম, সেটা এবার গুছিয়ে লিখে প্রমাণ করব। কিন্তু তার আগে একটা সহজ জিনিস লক্ষ্য করো। A যদি $m \times n$ সাইজের একটা matrix হয়, তবে $\dim(\mathscr{C}(A))$ এবং $\dim(\mathscr{R}(A))$ দুজনেই m আর n দুজনের থেকেই ছোটো হবে। কারণ $\mathscr{C}(A) \subseteq \mathbb{R}^m$, তাই $\dim(\mathscr{C}(A)) \leq m$ হতে বাধ্য। আবার A-র মধ্যে column আছেই মোটে n-খানা। তাদের span-ই হল $\mathscr{C}(A)$. তার মানে $\mathscr{C}(A)$-এর একটা generating set তো পেয়েই গেছি, যার সাইজ n. এদিকে basis হল একটা minimal generating set. তাই basis-এর সাইজ কোনোভাবেই n-এর চেয়ে বেশী হতে পারে না। সেই কারণে $\dim(\mathscr{C}(A)) \leq n$. একইরকম যুক্তি খাটবে $\mathscr{R}(A)$-র বেলাতেও।

> **THEOREM**
>
> For any matrix A
> $$\dim(\mathscr{C}(A)) = \dim(\mathscr{R}(A)).$$

Proof: প্রথমে সব কিছুর নাম দিয়ে নিই--

> Let A be $m \times n$.
> Let $r = \dim(\mathscr{C}(A))$.

এবার দুটো কেস হবে, প্রথমটা trivial–

> If $r = 0$, then $\mathscr{C}(A)$ is trivial vector space.
> $\therefore A = O$, and so $\mathscr{R}(A)$ is also trivial.
> $\therefore \dim(\mathscr{R}(A)) = 0 = \dim(\mathscr{C}(A))$, as required.

যুক্তিটা বুঝলে তো? Dimension শূন্য হয় কেবলমাত্র trivial vector space-এর বেলাতেই। আর trivial vector space-এর মধ্যে খালি $\vec{0}$-ই থাকে। তাই $\dim(\mathscr{C}(A))$ হওয়া মানে A-র প্রত্যেকটা column-ই আসলে $\vec{0}$. অর্থাৎ $A = O$. সুতরাং $\mathscr{R}(A)$-ও trivial vector space হতে বাধ্য!

এইবার বাকি কেসটা। আসল পরিশ্রমটা এখানেই হবে। এখানে মোক্ষম হাতিয়ার হবে contraction theorem. সেটা কী জিনিস ছিল মনে আছে তো? কিছু vector-এর একটা তালিকা দেওয়া থাকলে তার span যে একটা vector space হয় সেটা তো জানোই। Contraction theorem বলে যে, সেই তালিকার মধ্যেই এমন কিছু vector পাওয়া যাবে যারা মিলে এই vector space-টার একটা basis হবে। এখানে আমরা এই theorem-টা লাগাব A-র column-গুলোর উপরে।

> If $r > 0$, then by contraction theorem, we can find r columns of A that form a basis of $\mathscr{C}(A)$.

একটা উদাহরণ দেখলে হয়তো সুবিধা হবে। ধরো

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

এখানে দেখতেই পাচ্ছ যে, দ্বিতীয় column-টা আসলে প্রথম আর তৃতীয় column-এর linear combination. (ওদের যোগ করে $\frac{1}{2}$ দিয়ে গুণ করে দিলেই দ্বিতীয় column-টা পেয়ে যাবে)। কিন্তু প্রথম আর তৃতীয় column পরস্পরের সঙ্গে linearly independent. সুতরাং এখানে প্রথম আর তৃতীয় column মিলে $\mathscr{C}(A)$-র একটা basis হচ্ছে।

> We form a matrix $B_{m\times r}$ with these basis vectors as its column:
> $$B = \left[\begin{array}{ccc} \vec{b}_1 & \cdots & \vec{b}_r \end{array} \right].$$

আমাদের উদাহরণে B হবে
$$B = \left[\begin{array}{cc} 1 & 3 \\ 4 & 6 \\ 7 & 9 \end{array} \right].$$

> Then $\mathscr{C}(A) = \mathscr{C}(B)$, and hence $\mathscr{C}(A) \subseteq \mathscr{C}(B)$.
> So $A = BX$ for some matrix $X_{r\times n}$.

এইখানে আমরা একটু আগে শেখা theorem-টা লাগালাম। যদি বুঝতে অসুবিধা হয়, তবে আমাদের উদাহরণটা দেখে নাও। A-র প্রতিটা column-কেই B-এর column-গুলোর linear combination হিসেবে লেখা যায়। সেই linear combination-এর coefficient-গুলো জমিয়েই X-এর জন্ম। যদি A-র column-গুলোকে $\vec{a}_i$ আর B-এর column-গুলোকে $\vec{b}_i$ বলি তবে আমাদের উদাহরণে--

$$\begin{array}{rcl} \vec{a}_1 & = & 1\cdot\vec{b}_1 + 0\cdot\vec{b}_2, \\ \vec{a}_2 & = & \frac{1}{2}\cdot\vec{b}_1 + \frac{1}{2}\cdot\vec{b}_2, \\ \vec{a}_3 & = & 0\cdot\vec{b}_1 + 1\cdot\vec{b}_2. \end{array}$$

সুতরাং
$$\underbrace{\left[\begin{array}{ccc} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{array} \right]}_{A} = \underbrace{\left[\begin{array}{cc} \vec{b}_1 & \vec{b}_2 \end{array} \right]}_{B} \underbrace{\left[\begin{array}{ccc} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \end{array} \right]}_{X}.$$

এবার লক্ষ্য কর, X-কে বাঁদিক থেকে B দিয়ে গুণ করে A পাওয়া যাচ্ছে, সুতরাং--

> So $\mathscr{R}(A) \subseteq \mathscr{R}(X)$.
> $\therefore \dim(\mathscr{R}(A)) \leq \dim(\mathscr{R}(X)) \leq r$, since X has r rows.

এখানে $\dim(\mathscr{R}(X)) \leq r$ হচ্ছে, কারণ $X_{r\times n}$-এর row-সংখ্যা হল r, আর আমরা একটু আগেই দেখলাম যে, কোনো matrix-এর row space-এর dimension তার row আর column-এর সংখ্যার চেয়ে বেশী হতে পারে না।

> Thus we have shown that for any matrix A
> $$\dim(\mathscr{R}(A)) \leq \dim(\mathscr{C}(A)). \qquad (*)$$

এইবার একই কাজ করব A' নিয়ে--

> Applying the same result to A' in place of A we get
> $$\dim(\mathscr{R}(A')) \leq \dim(\mathscr{C}(A')).$$
> But $\mathscr{R}(A) = \mathscr{C}(A')$ and $\mathscr{C}(A) = \mathscr{R}(A')$. So
> $$\dim(\mathscr{C}(A)) \leq \dim(\mathscr{R}(A)). \qquad (**)$$

Hence from (*) and (**),

$$\dim(\mathscr{R}(A)) = \dim(\mathscr{C}(A)),$$

as required.

[Q.E.D]

এই যে দুটো dimension সমান হয়ে গেল, এই dimension-টাকে বলে A-র **rank.** এটা একটা অত্যন্ত গুরুত্বপূর্ণ ধারণা।

DEFINITION: Rank

For any matrix A the dimension of $\mathscr{C}(A)$ is called the **rank** of A. It is also the dimension of $\mathscr{R}(A)$. The rank of A is written as $r(A)$.

কেউ কেউ আবার $\dim(\mathscr{R}(A))$-কে A-র **row rank** এবং $\dim(\mathscr{C}(A))$-কে A-র **column rank** বলতে ভালোবাসেন। তবে যেহেতু দুটোই আসলে সমান তাই অধিকাংশ জায়গাতেই খালি rank শব্দটাই ব্যবহৃত হয়। কাল আর পরশু আমরা এ নিয়ে বিস্তারিত আলোচনা করব। আপাততঃ কিছু সহজ আলোচনা করে rank-এর সাথে বন্ধুত্ব করে নেওয়া যাক।

Exercise 1: 5×5 সাইজের এই দুটো matrix দ্যাখো--

$$A_1 = [\ \vec{e}_1 \quad \vec{0} \quad \vec{0} \quad \vec{0} \quad \vec{0}\ ], A_2 = [\ \vec{e}_1 \quad \vec{e}_2 \quad \vec{0} \quad \vec{0} \quad \vec{0}\ ].$$

এদের rank কী কী? এখানে $\vec{e}_i$ মানে এমন একটা 5-dimensional vector যার খালি i-th entry-টা 1, বাকিরা 0. ■

কোনো matrix-এর row rank বা column rank কখনোই তার row আর column-এর সংখ্যাকে ছাড়িয়ে যেতে পারে না। এটা আমরা আগেই আলোচনা করেছি। তা থেকেই আসে--

THEOREM

For any matrix A of size $m \times n$ we have

$$r(A) \leq \min\{m, n\}.$$

আচ্ছা, 0 থেকে $\min\{m, n\}$-এর মধ্যে যেকোনো integer-ই rank হতে পারে? মানে যদি তোমাকে আমি যা খুশি m, n দিই, আর কোনো একটা $r \in \{0, ..., \min\{m, n\}\}$ দিই, তবে কি তুমি সর্বদাই এমন $A_{m\times n}$ দিতে পারবে, যাতে $r(A) = r$ হয়? উত্তর হল--হ্যাঁ, পারবে। কেন পারবে সেইটা নিয়ে একটু গবেষণা করা যাক।

Exercise 2: চট্ করে 5×5 সাইজের একটা matrix দাও তো যার rank হবে 3.
HINT: আগের উদাহরণের অনুকরণে এগোও। ■

এই অংকটা করার পর নীচেরটা করা সহজ হওয়া উচিত।

Exercise 3: আগের অংকটার অনুকরণে পাঁচটা matrix দাও $A_0, A_1, ..., A_4$, যাদের প্রত্যেকেরই সাইজ 4×10, এবং rank যথাক্রমে $0, 1, ..., 4$. এবার যদি একই কাজ করতে বলতাম 10×4 সাইজ নিয়ে? ■

আমাদের গবেষণা শেষ হবে এই অংকটা দিয়ে--

Exercise 4: Let F be our field. Let $m, n \in \mathbb{N}$. Let k be any nonnegative integer $\leq \min\{m, n\}$. Show that there is an $m \times n$ matrix A such that $r(A) = k$. ■

DAY 17
$\mathscr{C}(A), \mathscr{R}(A)$ আর $\mathscr{N}(A)$ (part 2)

17.1 Rank বার করা

Rank বার করা মানে row space বা column space-র dimension বার করা। এর একটা কায়দা যে RREF দিয়ে, সে তো জানোই। তবে সহজ ক্ষেত্রে অনেক সময়ে চোখে দেখেই rank বলে দেওয়া যায়। যেমন matrix-টা যদি zero matrix হয় তবে তার rank হবে 0. যেসব matrix-এর rank হয় 1, তাদেরকেও চোখে দেখেই চেনা যায়, কারণ তাদের প্রতিটা row-ই একই nonzero vector-এর multiple হয়, যেমন এই matrix-টায়

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 10 & 20 & 30 & 40 \\ 100 & 200 & 300 & 400 \end{bmatrix}.$$

যদি অন্ততঃ দুটো row (বা দুটো column) পাও যারা পরস্পরের multiple নয়, অমনি বলতে পারো যে rank অন্ততঃ 2 হবেই। তবে চোখে দেখে এর বেশী এগোনো অনেক সময়েই কঠিন। কখনো আবার row-দের নিয়ে চিন্তা না করে column-দের নিয়ে চিন্তা করলে সুবিধা হয়--

Example 3: এই matrix-টার rank বার কর দেখি।

$$A = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 2 \\ 3 & 6 & 3 \end{bmatrix}.$$

SOLUTION: এখানে প্রথম আর তৃতীয় column সমান, আর দ্বিতীয় column-টা প্রথম column-এর multiple নয়। সুতরাং $\mathscr{C}(A)$-এর dimension হবে 2. মানে $r(A) = 2$. এখানে row নিয়ে চিন্তা করলে এত সহজে হত না। ■

Exercise 5: খালি চোখে দেখে নীচের matrix-গুলোর rank বার কর। চিন্তা করার সময়ে matrix-টার কথা ভুলে গিয়ে ওটাকে খালি কয়েকটা column-এর (বা row-এর) গুচ্ছ বলে ভাবলে সুবিধা হবে।

(1) $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$, **(2)** $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, **(3)** $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \end{bmatrix}$, **(4)** $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$.

(5) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

■

Exercise 6: যে matrix-এর সবগুলো entry-ই 0, তার rank-ও 0. এর উল্টোদিকটাও কি ঠিক, মানে $r(A) = 0$ হলে A-র সব entry-ই কি 0 হতে বাধ্য? ■

Exercise 7: Determine rank of row-space of

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

[2] **(2004.14cii)**

HINT:

এখানে ভাষাটা সঠিক নয়। হয় বলা উচিত matrix-টার rank, নয়তো row space-এর dimension. ■

এই গেল চোখে দেখে rank বার করার গল্প। যদি একটু বড় matrix হয়, তবে rank বার করতে RREF ব্যবহার করাই সহজ পথ। যদিও বর্তমান সিলেবাসে RREF নেই, কিন্তু তাও এই পদ্ধতির গুরুত্ব বিবেচনা করে, একটা উদাহরণ দেখাচ্ছি। সিলেবাসের বাইরে থেকে হাওয়া বইলে যাদের মাথা ধরে, তাদের মনে করিয়ে দিই যে, বিভিন্ন প্রতিযোগিতামূলক পরীক্ষায় MCQ-এর উত্তর বার করতে RREF কাজে লেগে যেতে পারে।

Example 4:

$$A = \begin{bmatrix} 1 & -1 & 2 & 0 & 4 \\ 2 & 2 & 1 & 5 & 2 \\ 1 & 3 & -1 & 0 & 3 \\ 1 & 7 & -4 & 1 & 1 \end{bmatrix}.$$

Transform A to row reduced echelon form and find its rank.[3+1] **(2007)**

SOLUTION:

$$\begin{bmatrix} \boxed{1} & -1 & 2 & 0 & 4 \\ 2 & 2 & 1 & 5 & 2 \\ 1 & 3 & -1 & 0 & 3 \\ 1 & 7 & -4 & 1 & 1 \end{bmatrix}$$

$$\xrightarrow[\substack{newR_3 = R_3 - R_1 \\ newR_4 = R_4 - R_1}]{newR_2 = R_2 - 2R_1} \begin{bmatrix} 1 & -1 & 2 & 0 & 4 \\ 0 & \boxed{4} & -3 & 5 & -6 \\ 0 & 4 & -3 & 0 & -1 \\ 0 & 8 & -6 & 1 & -3 \end{bmatrix}$$

$$\xrightarrow[\substack{newR_1 = R_1 + \frac{1}{4}R_2 \\ newR_3 = R_3 - R_2 \\ newR_4 = R_4 - 2R_2}]{newR_2 = R_2/4} \begin{bmatrix} 1 & 0 & \frac{5}{4} & \frac{5}{4} & \frac{5}{2} \\ 0 & 1 & -\frac{3}{4} & \frac{5}{4} & -\frac{3}{2} \\ 0 & 0 & 0 & \boxed{-5} & 5 \\ 0 & 0 & 0 & -9 & 9 \end{bmatrix}$$

pivot-টা এক লাফে এই অবস্থানে আসে নি।

প্রথমে $(3,3)$ অবস্থানে এসেছিল।

সেখানে nonzero আনতে ব্যর্থ হয়ে $(3,4)$ অবস্থানে এসেছে।

$$\xrightarrow[\substack{newR_1 = R_1 + \frac{1}{4}R_3 \\ newR_2 = R_2 + \frac{1}{4}R_3 \\ newR_4 = R_1 - \frac{9}{5}R_3}]{newR_3 = -R_3/5} \begin{bmatrix} 1 & 0 & \frac{5}{4} & 0 & \frac{15}{4} \\ 0 & 1 & -\frac{3}{4} & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & \boxed{0} \end{bmatrix}$$

> This matrix has 3 nonzero rows. So A has rank 3.

■

যদি নিজে নিজে এরকম একটা অংক কষে দেখার জন্য হাত নিশপিশ করে, তবে নীচের অংকটা করে দেখতে পারো।

Exercise 8: Reduce the matrix A to a row-reduced-echelon form and hence find its rank, where

$$A = \begin{bmatrix} 0 & 1 & -3 & -1 \\ 1 & 0 & 1 & 1 \\ 3 & 1 & 0 & 2 \\ 1 & 1 & 2 & 0 \end{bmatrix}.$$

[5] **(2011.10a)** ■

এতক্ষণ সবগুলো অংকতেই কাজ করছিলাম row নিয়ে, এবার column দিয়ে একটা একইরকম অংক দেখি।

Example 5: Find the column space and the column rank of the matrix

$$A = \begin{bmatrix} 2 & 1 & 4 & 3 \\ 3 & 2 & 6 & 9 \\ 1 & 1 & 2 & 6 \end{bmatrix}$$

(2009.12a)

SOLUTION: এটা ফের চোখে দেখেই করে ফেলার অংক। এখানে কাজ করতে হবে column-গুলো নিয়ে।

> Let the columns be $\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4$.

তৃতীয় column-টা দেখাই যাচ্ছে প্রথমটার 2-গুণ। আর চতুর্থ column-টা হল প্রথম column-এর 3-গুণ।

> Here $\vec{a}_3 = 2\vec{a}_1$ and $\vec{a}_4 = 3\vec{a}_1$.

সুতরাং খালি $\vec{a}_1$-কে রাখলেই হবে, $\vec{a}_3$ আর $\vec{a}_4$-কে ফেলে দিলেই হবে। এবার দেখি $\vec{a}_2$-কেও বিদায় করা যায় কিনা।

> But $\vec{a}_2$ is not a multiple of $\vec{a}_1$.

সুতরাং $\vec{a}_2$-কে রাখতে হবে--

> Thus $\mathscr{C}(A)$ is the space with basis $\{\vec{a}_1, \vec{a}_2\}$.

এই basis-এর সাইজ হল 2, তাই--

> The column rank of A is 2.

যদি এভাবে চোখে দেখে না করা যেত, তবে RREF-এর শরণাপন্ন হতে হত। যেহেতু এখানে $\mathscr{C}(A)$ বার করতে বলেছিল, তাই আমরা A'-এর RREF বার করতাম। ■

চোখে দেখেই বার করি, আর row/column echelon form দিয়েই বার করি, rank বার করার জন্য matrix-এর কোথায় কী সংখ্যা আছে জানা থাকলে সুবিধা হয়। কিন্তু যদি কোনো matrix-এর entry-গুলো সংখ্যার বদলে a, b, c ইত্যাদি অক্ষর দিয়ে দেওয়া থাকে তবে মুস্কিল। ওরা zero নাকি nonzero না জানা থাকলে ওদের দিয়ে ভাগ করা যায় না, তাই linear combination বানাতে অসুবিধা হয়। নীচের অংকটা এরকম একটা উদাহরণ।

Example 6: If a, b, c are unequal real numbers then find the rank of the matrix

$$\begin{bmatrix} 0 & a & b \\ -a & 0 & -c \\ -b & c & 0 \end{bmatrix}.$$

[2] **(2007.12ci)**

SOLUTION: প্রথম row-টা শুরু হয়েছে 0 আর a দিয়ে, কিন্তু দ্বিতীয় row-এর গোড়ায় আছে $-a$ আর 0. যদি $a \neq 0$ হয়, তবে তৎক্ষণাৎ বলে দেওয়া যায় যে, ওই দুটো row হবে linearly independent. সমস্যা হল, $a \neq 0$ বলে দেয় নি। তাই দুটো কেসে ভেঙে এগোলে সুবিধা হবে।

> Case 1 $a = 0$: Here the matrix is
>
> $$\begin{bmatrix} 0 & 0 & b \\ 0 & 0 & -c \\ -b & c & 0 \end{bmatrix}.$$
>
> The first two rows are both nonzero multiples of $(0, 0, 1)$.

এখানে "nonzero" বলতে পারলাম, কারণ a, b, c সকলে আলাদা বলে দিয়েছে, আর $a = 0$ ধরে নিয়েছি। তাই $b, c \neq 0$. তার মানে ওরা দুজনেই একই লাইন বরাবর আছে, তাই rank-এ ওদের অবদান হবে 1.

> Also since $b, c \neq a = 0$, the third row is not a multiple of $(0, 0, 1)$.

যেহেতু তৃতীয় row-টা সেই লাইনের থেকে বাইরে মুখ বাড়িয়েছে, তাই ও rank-এ আরো 1 যোগ করবে। সুতরাং rank-টা দাঁড়াচ্ছে $1 + 1 = 2$.

> So the rank is 2.

এবার দেখি $a \neq 0$ হলে কী হয়। এখানে প্রথম দুটো row তো linearly independent হয়েই আছে। খালি দেখতে হবে তৃতীয় row-টাকে ওদের linear combination হিসেবে লেখা যাচ্ছে কিনা। যদি যায়, তবে rank=2 হবে, নইলে rank=3 হবে। তার মানে আমরা চাইছি--

> Case 2 $a \neq 0$:
>
> Here the first two rows are obviously linearly independent.
>
> Want to check if $(b, c, 0) = p(0, a, b) + q(-a, 0, -c)$ for some p, q,
>
> ie, $b = -aq,\ c = pa$ and $0 = pb - qc$.
>
> From the first two equations $q = -\frac{b}{a}$ and $p = \frac{c}{a}$.

এবার দুরুদুরু বক্ষে তৃতীয় equation-টায় এদেরকে বসিয়ে দেখতে হবে মেলে কিনা--

> Putting these in the last equation we get $0 = \frac{c}{a}b - \frac{b}{a}c$, which is indeed true.
> Thus the third row is in the span of the first two. Hence the rank is 2.

একটা ছোট্টো উপসংহার করে দিই--

> Thus in both the cases the rank is 2.

■

17.2 $\mathscr{N}(A)$ আর $\mathscr{R}(A)$-এর সম্পর্ক

আমরা column space আর row space-এর মধ্যে দুটো সম্পর্ক শিখেছি--এক, $\mathscr{C}(A)$ আর $\mathscr{R}(A)$-র dimension সবসময় সমান হয়, আর দুই, $\mathscr{C}(A') = \mathscr{R}(A)$ হয়। এবার $\mathscr{R}(A)$ এবং $\mathscr{N}(A)$-র dimension-এর মধ্যে একটা গুরুত্বপূর্ণ সম্পর্ক শিখতে চলেছি।
চট্ করে মনে করিয়ে দিই $\mathscr{N}(A)$ মানে কী ছিল। এটা হল সেই সব vector-এর set, যাদেরকে বাঁদিক থেকে A দিয়ে গুণ করলে $\vec{0}$ পাওয়া যায়, মানে যদি field-টা F হয়, আর A হয় $m \times n$ সাইজের একটা matrix, তবে--

$$\mathscr{N}(A) = \{\vec{x} \in F^n \ : \ A\vec{x} = \vec{0}\}.$$

এর একটা নাম A-র **nullspace**, আরেকটা নাম হল A-র **kernel.** এটা যে F^n-এর একটা subset শুধু তাই নয়, একটা subspace-ও বটে। এটা আমরা আগেই বলেছিলাম। যেটা বলি নি সেটা হল এই যে, এর dimension-এরও একটা বিশেষ নাম আছে।

> DEFINITION: **Nullity**
>
> For any matrix A we call $\dim(\mathscr{N}(A))$ its **nullity**.

যদি A হয় $m \times n$ সাইজের, তবে $\mathscr{N}(A) \subseteq F^n$. তাই $0 \leq \text{nullity}(A) \leq n$ হবে। মানে nullity কখনো column-এর সংখ্যার বেশী হতে পারে না। এটা বোঝা সহজ। কিন্তু nullity-র সঙ্গে rank-এর একটা আশ্চর্য সম্পর্ক আছে, যেটা বলে না দিলে কল্পনা করাও কঠিন--

> THEOREM
>
> For any matrix A over any field F
>
> $$r(A) + \text{nullity}(A) = \text{number of columns of } A.$$

এই theorem-এর প্রমাণটা একটু কৌশলের। সেটা শেখার আগে প্রথমে theorem-টা হজম করা দরকার, তার জন্য একগুচ্ছ অংক দিলাম নীচে। প্রমাণে হাত দেব তার পরে।

Exercise 9: যদি $r(A_{4\times 5}) = 2$ হয়, তবে $\text{nullity}(A)$ কত? ■

Exercise 10: একটা $m \times n$ সাইজের matrix-এর rank আর nullity সমান। তবে নীচের কোন্ কথাটা সঠিক?

(A) m-টা even (জোড়) হতে বাধ্য।

(B) m-টা odd (বিজোড়) হতে বাধ্য।

(C) n-টা even হতে বাধ্য।

(D) n-টা odd হতে বাধ্য।

■

Exercise 11: ধরো আমাদের field-টা হল F, আর A হল তার উপরে $m \times n$ সাইজের একটা matrix. তবে $\mathscr{C}(A)$, $\mathscr{R}(A)$ এবং $\mathscr{N}(A)$ এরা সকলেই আছে হয় F^m-এর মধ্যে আর নয়তো F^m-এর মধ্যে। কে কোথায় আছে? ■

Exercise 12: একটা 5×7 সাইজের matrix আছে। বলা আছে যে row-গুলো মিলে একটা linearly independent set হয়। Matrix-টার nullity কী হবে? ■

Exercise 13: $\text{nullity}(A_{m\times n}) = 0$. যদি $m \neq n$ হয়, তবে নীচের কোন কথাটা সত্যি হতে বাধ্য?

1. $m < n$
2. $m > n$
3. জোর দিয়ে কোনোটাই বলা যায় না।

■

Exercise 14: একটা square matrix দিলাম $n \times n$, যার column-গুলো মিলে একটা linearly independent set হয়। তবে নীচের কোন বাক্যগুলো সঠিক?

1. $\mathscr{C}(A) = \mathbb{R}^n$
2. $\mathscr{R}(A) = \mathbb{R}^n$
3. $\mathscr{N}(A) = \mathbb{R}^n$
4. row-গুলো মিলেও একটা linearly independent set হবে।

■

Exercise 15: A একটা 10×20 সাইজের matrix. যদি $\text{nullity}(A') = 1$ হয়, তবে $\text{nullity}(A)$ কত হবে? ■

Exercise 16: একটা matrix আছে A এমন যে $\text{nullity}(A) = \text{nullity}(A')$. তাহলে নীচের কোন কথাটা ঠিক হতে বাধ্য?

1. $\text{nullity}(A) = 0$.

2. A অবশ্যই square.

3. $A = A'$.

4. এর কোনোটাই নয়।

■

যদি A হয় $m \times n$ সাইজের, তবে $\mathscr{R}(A)$ আর $\mathscr{N}(A)$ দুজনেই হল F^n-এর দুটো subspace. আবার ওদের dimension যোগ করলে F^n-এর dimension পাওয়া যাচ্ছে। এইটা দেখে সন্দেহ জাগতে পারে যে, $\mathscr{R}(A)$ আর $\mathscr{N}(A)$ বুঝি পরস্পরের direct complement, কারণ direct complement-দের dimension যোগ করলে যে পুরো vector space-এর dimension পাওয়া যায়, সেটা আমরা আগেই দেখেছি। কিন্তু না, $\mathscr{R}(A)$ আর $\mathscr{N}(A)$ পরস্পরের direct complement নাই হতে পারে। এরকম একটা উদাহরণ আছে নীচের অংকটায়।

Exercise 17: ধরো আমাদের field হল $\mathbb{C}$. এই matrix-টা নাও, $A = \begin{bmatrix} 1 & i \\ 1 & i \end{bmatrix}$. দেখাও যে $\mathscr{R}(A) \cap \mathscr{N}(A) \neq \{\vec{0}\}$. তাতেই প্রমাণ হয়ে যাচ্ছে যে, $\mathscr{R}(A)$ আর $\mathscr{N}(A)$ পরস্পরের direct complement নয়। ■

তবে যদি field-টা $\mathbb{R}$ হয়, তবে $\mathscr{R}(A)$ আর $\mathscr{N}(A)$ সত্যই পরস্পরের direct complement হয়। শুধু তাই, ওরা খুব বিশেষ এক ধরণের direct complement হয়, যাকে বলে orthogonal complement. এদের কথা আমরা এই বইয়ের অষ্টম অধ্যায়ে গিয়ে শিখব।

17.2.1 প্রমাণ

Theorem-টা আমরা প্রমাণ করব সরাসরি dimension-এর সংজ্ঞা ব্যবহার করে। ধাপগুলো জানা না থাকলে ভেবে বার করা একটু কঠিন। প্রমাণটা পড়ার আগে একটু সাবধান করে দিই। rank আর nullity যোগ করলে যে column-এর মোট সংখ্যা পাওয়া যায়, এর মত সহজ কথা আর কিছুই হতে পারে না, এবং এই কথাটুকু মনে রাখাই theorem-টাকে প্রয়োগ করার পক্ষে যথেষ্ট। প্রমাণের জটিল ধাপগুলো সেখানে কাজে লাগবে না। সুতরাং প্রমাণটার দৈর্ঘ্য দেখে যেন মুষড়ে যেও না। বাগান তৈরী করা যতই দুরূহ হোক, বাগানের সৌন্দর্য উপভোগ করার পথে তা যেন অন্তরায় হয়ে না দাঁড়ায়!

শুরুতেই একটা জিনিস মনে করিয়ে দিই--যদি $A_{m\times n}$ কোনো একটা matrix হয়, এবং $\{\vec{e}_1, ..., \vec{e}_n\}$ হয় আমাদের পরিচিত basis-টা, তবে $A\vec{e}_i$ হবে A-র i-th column-টা। যেমন

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}.$$

এইবার প্রমাণটা শুরু করা যাক।

> Proof of the theorem: Let $A_{m\times n}$ be any matrix over a field F.
> Let $r = r(A)$ and $k = \text{nullity}(A)$.

আমরা দেখাব যে $r + k = n$. তার জন্য আমরা F^n-এর একটা basis বার করে দেখাব যার সাইজ $r + k$. ব্যস্, তাহলেই হয়ে যাবে, কারণ F^n-এর dimension তো n. এই কাজটা আমরা ধাপে ধাপে করব। প্রথমে F^n-এর একটা subset বানাব B. তার পর দেখাব যে B-এর সাইজ হল $r + k$. তারপর দেখাব যে B হল F^n-এর একটা basis.

প্রথম ধাপে set-টা বানানো যাক--

> Step 1: W.l.g., suppose that the first r columns of A form a basis of $\mathscr{C}(A)$.

এই একটা বাক্যের মধ্যে অনেক কথা লুকিয়ে আছে। $\mathscr{C}(A)$ হল A-র column-গুলো দিয়ে generated. মানে column-গুলো মিলে $\mathscr{C}(A)$-র একটা generating set হয়। Contraction theorem বলে যে এর একটা subset পাব, যেটা $\mathscr{C}(A)$-র basis হবে, মানে এমন r-খানা column আছে, যারা মিলে $\mathscr{C}(A)$-এর একটা basis হবে। কাজের সুবিধার জন্য ধরে নিয়েছি এরা হল প্রথম r-খানা column. প্রথম r-খানা না হলেও আপত্তি ছিল না, খালি সেক্ষেত্রে বলতে হত $j_1, ..., j_r$ নম্বর column-গুলো একটা basis তৈরী করে, তাতে notation-এর জটিলতা খামোখা বেড়ে যেত। সেই জন্যই without loss of generality (wlg) লিখে ধরে নিয়েছি যে, প্রথম r-খানা column মিলে $\mathscr{C}(A)$-র একটা basis হয়।

> Let these columns be $\vec{a}_1, ..., \vec{a}_r$.
>
> Let $\vec{e}_i \in F^n$ be the vector with all entries 0, except the i-th entry, which is 1.
>
> Let $\{\vec{n}_1, ..., \vec{n}_k\} \subseteq F^n$ be a basis of $\mathscr{N}(A)$.
>
> Let $B = \{\vec{e}_1, ..., \vec{e}_r, \vec{n}_1, ..., \vec{n}_k\}$.

এই হল আমাদের set-টা। এবার দেখাব যে এর সাইজ $r + k$. মনে হতে পারে যে, সেটা বলাই বাহুল্য, কারণ এর মধ্যে r-খানা $\vec{e}_i$ আর k-খানা $\vec{n}_j$ আছে। কিন্তু আমরা এখনও জানি না যে, এরা সবাই আলাদা আলাদা কিনা। যেমন $\{1, 2, 3, 2\}$-এর সাইজ কিন্তু 4 নয়, ওর সাইজ আসলে 3, যেহেতু 2 সংখ্যাটা দুবার লেখা হলেও আলাদা করে গোনা হবে না।

> Step 2: Shall show that B is a set of size $r + k$. By definition $\vec{e}_i$'s are all distinct.
>
> Since $\{\vec{n}_1, ..., \vec{n}_k\}$ is a linearly independent set, so $\vec{n}_j$'s are all distinct.
>
> For $i = 1, ..., r$, we have $A\vec{e}_i = \vec{a}_i \neq \vec{0}$, since $\{\vec{a}_1, ..., \vec{a}_r\}$ is linearly independent.

এখানে আমরা সেই কথাটা ব্যবহার করলাম, $A\vec{e}_i$ হল A-র i-th column.

> So $\vec{e}_1, ..., \vec{e}_r \notin \mathscr{N}(A)$.
>
> Hence $\vec{e}_i$'s are distinct from $\vec{n}_j$'s.
>
> So $|B| = r + k$.

এইবার B-কে basis দেখানো শুরু হবে। প্রথমে linearly independent দেখাই--

> Step 3: Shall show that B is a linearly independent set.

সেই পরিচিত কায়দা, যেকোনো একটা linear combination-কে $\vec{0}$ ধরে নিয়ে সবগুলো coefficient-কে 0 দেখানো--

> Let $c_1, ..., c_r, d_1, ..., d_k \in F$ be such that
>
> $$\sum_1^r c_i\vec{e}_i + \sum_1^k d_j\vec{n}_j = \vec{0}.$$
>
> Shall show that $\forall i \quad c_i = 0$ and $\forall j \quad d_j = 0$.

Multiplying by A from left,

$$\begin{aligned} & A\left[\sum_1^r c_i\vec{e}_i + \sum_1^k d_j\vec{n}_j\right] = A\vec{0}, \\ \text{or,} \quad & \sum_1^r c_iA\vec{e}_i + \sum_1^k d_jA\vec{n}_j = \vec{0}, \\ \text{or,} \quad & \sum_1^r c_iA\vec{e}_i + \sum_1^k d_jA\vec{n}_j = \vec{0}, \\ \text{or,} \quad & \sum_1^r c_i\vec{a}_i = \vec{0}, \quad [\because \forall j \quad A\vec{n}_j = \vec{0}] \\ \text{or,} \quad & c_1 = \cdots = c_r = 0, \end{aligned}$$

Since $\{\vec{a}_1, ..., \vec{a}_r\}$ is linearly independent.

So $\sum_1^k d_j\vec{n}_j = \vec{0}$, and hence $d_1 = \cdots = d_k = 0$, since $\{\vec{n}_1, ..., \vec{n}_k\}$ is linearly independent.

ব্যস্, linearly independent দেখানো শেষ হল। এবার generating set দেখাব--

Step 4: Shall show that $\text{span}(B) = F^n$.

Since $B \subseteq F^n$, so $span(B) \subseteq F^n$.

তার মানে আমাদের খালি দেখাতে হবে উল্টোদিকটা, অর্থাৎ F^n-এর যেকোনো vector-কেই B-এর linear combination হিসেবে লেখা যায়।

Let $\vec{v} \in F^n$.
Then $A\vec{v} \in \mathscr{C}(A)$,

কারণ A-কে যেকোনো vector দিয়ে ডান দিকে গুণ করলে গুণফলটা $\mathscr{C}(A)$-র মধ্যে থাকে। এদিকে $\mathscr{C}(A)$-র একটা basis হল $\{\vec{a}_1, ..., \vec{a}_r\}$. তাই--

and so $A\vec{v} = p_1\vec{a}_1 + \cdots + p_r\vec{a}_r$, for some $p_1, .., p_r \in F$.
Let $\vec{w} = \vec{v} - \sum_1^r p_i\vec{e}_i$.
Then $\vec{w} \in \mathscr{N}(A)$.

⟦ Because:

$$\begin{aligned} A\vec{w} &= A\vec{v} - \sum_1^r p_iA\vec{e}_i \\ &= A\vec{v} - \sum_1^r p_i\vec{a}_i = \vec{0}. \end{aligned}$$

]]

Hence $\vec{w} = q_1\vec{n}_1 + \cdots + q_k\vec{n}_k$ for some $q_1, ..., q_k \in F$.
Thus $\vec{v} = \sum_1^r p_i\vec{e}_i + \sum_1^k q_j\vec{n}_j$.
Hence $F^n \subseteq \text{span}(B)$.
But $\text{span}(B) \subseteq F^n$.
So $\text{span}(B) = F^n$.

ব্যস্, প্রমাণ বলতে গেলে শেষ। খালি উপসংহারটুকু বাকি--

Step 5: Thus B is a basis of F^n. Hence $|B| = \dim(F^n) = n$.
Thus $r + k = n$, as required.

DAY 18
Rank নিয়ে আরো আলোচনা

Rank হল linear algebra-র অন্যতম গুরুত্বপূর্ণ ধারণা। এর সবচেয়ে বড় গুণ বোধ করি এই যে, এটা কোনো matrix-এর row space এবং column space-এর মধ্যে একটা যোগসূত্র রচনা করে-- এই দুটো space-এরই dimension একই হয়, এবং সেটাই হল rank. এর ফলে কোনো matrix-এর column সম্পর্কে কিছু তথ্য দেওয়া থাকলে আমরা অনেক সময়ে row-গুলো সম্বন্ধে তথ্য পেতে পারি (বা row-গুলো থেকে column-গুলো সম্বন্ধে)। এরকম একটা ধর্মের আলোচনা দিয়েই আজকের ক্লাস শুরু করা যাক।

18.1 Full column/row rank matrix

ধরো $A_{m \times n}$ একটা matrix. যদি এর column-গুলো সব linearly independent হয়, তবে A-কে বলে একটা **full column rank** matrix. একইভাবে, যদি row-গুলো linearly independent হয়, তবে বলে **full row rank** matrix.

DEFINITION: **Full column/row rank matrix**

A matrix with linearly independent columns is called a **full column rank** matrix.
A matrix with linearly independent rows is called a **full row rank** matrix.

কয়েকটা উদাহরণ দেখে না নিলে কোনো সংজ্ঞাই হজম হতে চায় না।

Example 7: নীচে কয়েকটা matrix দেওয়া আছে। এদের মধ্যে কে বা কারা full column rank, আর কারাই বা full row rank**?**

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{bmatrix},\ B = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},\ C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},\ D = \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ 5 & 10 \end{bmatrix}.$$

SOLUTION:

- A-র column space আছে $\mathbb{R}^2$-এর মধ্যে। Column-এর সংখ্যা 3. যেহেতু $\mathbb{R}^2$-এর dimension হল মোটে 2, তাই তিনটে column কোনোভাবেই linearly independent হতে পারে না। তাই A মোটেই full column rank নয়। কিন্তু ওর row দুটো কেউ কারো multiple নয়, তাই row-গুলো linearly independent. সুতরাং A হল full row rank.

- B-এর আছে খালি একটাই column, এবং সেটা nonzero. তাই ওই একখানা column একাই একটা linearly independent set তৈরী করে। অতএব B অবশ্যই full column rank. কিন্তু B একেবারেই full row rank নয়, কারণ দ্বিতীয় row-টাই দেখতে পাচ্ছি প্রথম row-এর multiple (দ্বিগুণ)।

- C-টা একই সঙ্গে full column rank এবং full row rank.

- যেহেতু $\mathbb{R}^2$-এর মধ্যে তিনটে vector কখনও linearly independent হতে পারে না, তাই D-এর পক্ষে full row rank হওয়া সম্ভব নয়। এখানে D-টা full column rank-ও নয়, কারণ ওর দ্বিতীয় column-টা হল প্রথমটার multiple.

■

Exercise 18: নীচের matrix-গুলোর মধ্যে কে বা কারা full column rank, আর কারাই বা full row rank?

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 4 & 7 \end{bmatrix},\ B = \begin{bmatrix} 0 & 0 & 3 \end{bmatrix},\ C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},\ D = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}.$$

■

Exercise 19: একটা $m \times n$ সাইজের matrix যদি full column rank হয়, তবে তার rank কত হবে? যদি full row rank হত, তবে? ■

Exercise 20: এমন একটা 3×5 সাইজের matrix দাও দেখি, যেটা একাধারে full row rank এবং full column rank. ■

সংজ্ঞাটা আশা করি এবার মাথায় ঢুকেছে। এবার দ্বিতীয় অধ্যায় থেকে কয়েকটা কথা মনে করিয়ে দিই। যদি $\vec{v}_1, ..., \vec{v}_k \in F^d$ কিছু vector হয়, আর $S = \text{span}\{\vec{v}_1, ..., \vec{v}_k\}$ হয়, তবে--

- $\dim(S) \leq k$ তো হবেই।
- আর $\dim(S) = k$ হওয়া এবং $\{\vec{v}_1, ..., \vec{v}_k\}$-এর linearly independent হওয়া সমার্থক।
- অপরপক্ষে, যদি $\dim(S) = d$ হয়, তবে $S = F^d$ হবে, এবং সেক্ষেত্রে $\{\vec{v}_1, ..., \vec{v}_k\}$ হবে F^d-র একটা generating set.

Exercise 21: A একটা 10×20 সাইজের real matrix (মানে এখানে field-টা $\mathbb{R}$). যদি $r(A) = 10$ হয়, তবে নীচের কোন কোন বাক্য সঠিক?

1. A-র column-গুলো মিলে একটা linearly independent set হয়,
2. A-র row-গুলো মিলে একটা linearly independent set হয়,

3. $\mathscr{C}(A) = \mathbb{R}^{10}$,
4. $\mathscr{C}(A) = \mathbb{R}^{20}$,
5. $\mathscr{R}(A) = \mathbb{R}^{10}$,
6. $\mathscr{R}(A) = \mathbb{R}^{20}$.

■

Exercise 22: একটা $m \times n$ সাইজের matrix-এর row-গুলো মিলে একটা linearly independent set হয়। এ থেকে কি জোর দিয়ে বলা যায় যে column-গুলো মিলে $\mathbb{R}^m$-এর একটা generating set হবেই? ■

Exercise 23: একটা matrix আছে, যার row-এর সংখ্যা 5. যদি তার column-গুলো সবাই linearly independent হয়, তবে সবচেয়ে বেশী কতগুলো column থাকতে পারে?
HINT: ধরো n-খানা column আছে। তারা সবাই linearly independent হলে rank কত হবে? মনে রেখো rank কখনোই row-এর সংখ্যাকে ছাড়িয়ে যেতে পারে না। ■

এই অংকগুলো থেকে যে শিক্ষাটা পাওয়া গেল, সেটা থেকেই নীচের theorem-টার জন্ম--

THEOREM

Let A be any $m \times n$ matrix over some field F. Then

1. the columns of A are linearly independent iff $\mathscr{R}(A) = F^n$.
2. the rows of A are linearly independent iff $\mathscr{C}(A) = F^m$.

প্রমাণ অতি সহজ, খালি প্রথম অংশটার প্রমাণ করছি। অন্য অংশের প্রমাণটা একইরকম।

Proof:

(1) The n columns of A are linearly independent

$$\begin{aligned} &\iff \dim(\mathscr{C}(A)) = n \\ &\iff r(A) = n \\ &\iff \dim(\mathscr{R}(A)) = n \\ &\iff \mathscr{R}(A) = F^n \quad [\because \mathscr{R}(A) \subseteq F^n] \end{aligned}$$

[Q.E.D]

এই theorem-টা থেকে একটা অপ্রত্যাশিত জিনিস আসে। বলো তো, $\mathscr{C}(I_m)$ কী? এখানে I_m মানে হল $m \times m$ সাইজের identity matrix. উত্তর হল $\mathscr{C}(I_m) = F^m$. সুতরাং A যদি full row rank হয়, তবে theorem-টা বলছে যে, $\mathscr{C}(A) = \mathscr{C}(I_m)$ হবে। তার মানে $\mathscr{C}(I_n) \subseteq \mathscr{C}(A)$ বলা যায় (ঠিক যেমন $2 \leq 2$ লেখা যায়)। অতএব I_m-কে নিশ্চয়ই AM আকারে লেখা যাবে, যেখানে M কোনো একটা matrix. এটা কেন হল, বুঝলে তো? যদি $\mathscr{C}(B) \leq \mathscr{C}(A)$ হয়,

তবে B-কে "$A\times$ কোনো matrix" আকারে লেখা যায়। একইরকম যুক্তি লাগানো যাবে full column rank-এর বেলাতেও। এ থেকেই আসবে--

THEOREM

- $A_{m\times n}$ is full column rank if and only if $NA = I_n$ for some $N_{n\times m}$. Any such N is called a **left inverse** of A.
- $A_{m\times n}$ is full row rank if and only if $AM = I_m$ for some $M_{n\times m}$. Any such M is called a **right inverse** of A.
- $A_{m\times n}$ is both full row rank and full column rank if and only if $m = n$ and $AB = BA = I$ for some B. This B is called the **inverse** of A.

এই theorem-টার বিষয়ে কিছু কথা--

- আমরা প্রথম অধ্যায়েই শিখেছিলাম যে, কোনো matrix-এর inverse থাকলে, তাকে বলে nonsingular. সুতরাং এখন তাদের সম্বন্ধে একটা নতুন তথ্য পেলে-- একটা matrix-এর nonsingular হওয়া মানে ওর একই সঙ্গে full column rank এবং full row rank হওয়া।
- লক্ষ্য করো, left আর right inverse-এর বেলায় আগে "a" লিখেছি, কিন্তু inverse-এর বেলায় লিখেছি "the". তার কারণ left বা right inverse একটা matrix-এর একাধিক থাকতেই পারে। কিন্তু inverse যদি থাকে তো ঠিক একটাই থাকবে। এই uniqueness-এর প্রমাণটা বেশ সহজ-- ধরো B আর C দুজনেই A-র inverse হবার দাবীদার। তবে
$$B = BI = B(AC) = (BA)C = IC = C.$$
- Full column rank-দের left inverse থাকে, আর full row rank-দের থাকে right inverse. এইটা অনেক সময়ে গুলিয়ে যেতে চায়। তাই এইভাবে মনে রাখতে পারো--coLumn-এও L আছে, Left-এও L আছে। আবার Row-তেও R আছে, Right-এও তাই।

Exercise 24: Theorem-টা গুছিয়ে লিখে প্রমাণ করো দেখি। সাবধান, এখানে কিন্তু "if and only if" আছে! ■

কোনো matrix যদি full column rank হয়, তবে তার rank হবে column-এর সংখ্যার সমান। যদি full row rank হত, তবে rank হত row-এর সংখ্যার সমান। এবং এর উল্টোকথাটাও সত্য। এইটা একটা theorem-এর আকারে লিখে রাখা যাক।

THEOREM

Let A be any matrix of size $m \times n$. Then

- A is full column rank if and only if $r(A) = n$,
- A is full row rank if and only if $r(A) = m$,
- A is nonsingular if and only if $r(A) = m = n$.

Example 8: What will be rank of an orthogonal matrix of order 4?[1] **(2007.9fp2)**

SOLUTION: Orthogonal matrix-রা সর্বদা nonsingular হয়, কারণ তার transpose-ই হয় তার inverse. তাই rank-টা হবে 4.

> An orthogonal matrix is nonsingular. So its rank must equal its order. So $r(A) = 4$.

■

--- এবার একটা ছোট্টো জিনিস--

> **THEOREM**
>
> A matrix has rank zero if and only if all its entries are 0.

এটা প্রমাণ করা এতই সহজ যে প্রমাণটা আর করে দিলাম না।

18.2 Transpose

আমরা দেখেছি যে rank-কে দুভাবে দেখা যায়--$\dim(\mathscr{C}(A))$ হিসেবে আর $\dim(\mathscr{R}(A))$ হিসেবে। যেহেতু একটা matrix-কে transpose করলেই তার row-গুলো column, আর column-গুলো row হয়ে যায়, তাই বুঝতেই পারছ যে--

> **THEOREM**
>
> For any matrix A, we have $r(A) = r(A')$.

Exercise 25: একটা full column rank matrix-এর transpose করলে কি একটা full row rank matrix পাওয়া যাবেই? ■

18.3 Submatrix

কোনো matrix-এর কিছু row আর column যদি বাদ দিয়ে দেওয়া যায় তবে যেটুকু পড়ে থাকে, তাকে বলে মূল matrix-টার একটা submatrix.

Example 9: যদি $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$ হয়, তবে প্রথম এবং তৃতীয় column এবং দ্বিতীয় row-টা বাদ দিলে কী submatrix পাবে?

SOLUTION: Submatrix-টা হবে $\begin{bmatrix} 2 & 4 \\ 10 & 12 \end{bmatrix}$. ■

সব সময়ে জানবে যে, submatrix-এর rank কখনও পুরো matrix-টার rank-এর চেয়ে বেশী হতে পারে না--

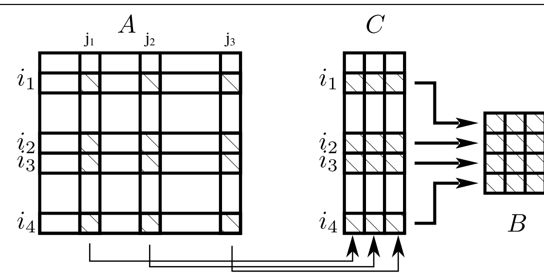


Fig 1

> THEOREM
>
> Let B be any submatrix of A. Then we must have $r(B) \leq r(A)$.

Proof:

> Let B consist of rows $i_1, ..., i_p$ and columns $j_1, ..., j_q$ of A.

তার মানে A থেকে যদি আমরা $i_1, ..., i_p$ নম্বর row-গুলো বাদে বাকি row-গুলোকে ফেলে দিই, এবং $j_1, ..., j_q$ ছাড়া বাকি column-দেরও ফেলে দিই, তবে যা পড়ে থাকবে সেটাই হল B.
এই ফেলে দেওয়ার কাজটাই আমরা দুই ধাপে ভেঙে করব। প্রথমে $j_1, ..., j_q$ বাদে বাকি column-দের ফেলে দেব, কিন্তু row রাখব সবগুলোই। এভাবে যে matrix-টা পাব তাকে বলি C. Fig 1 দ্যাখো।

> Let C be the submatrix of A consisting of columns $j_1, ..., j_q$, and all the rows.
>
> Then the set of columns of C is a subset of the set of columns of A.
>
> So $\mathscr{C}(C) \subseteq \mathscr{C}(A)$.
>
> So $r(C) \leq r(A)$.

এইবার C-এর থেকে $i_1, ..., i_p$ ছাড়া বাকি row-দের মুছে ফেলে B-তে পৌঁছব।

> Again the rows of B form a subset of the rows of C.
>
> $\therefore \mathscr{R}(B) \subseteq \mathscr{R}(C)$.
>
> $\therefore r(B) \leq r(C)$.
>
> Thus we have
>
> $$r(B) \leq r(C) \leq r(A),$$
>
> as required.

[Q.E.D]

18.4 গুণের প্রভাব

এবার rank নিয়ে আরেকটা গুরুত্বপূর্ণ theorem দিয়ে আলোচনা করব--গুণ করার ফলে rank কখনও বেড়ে যেতে পারে না। যদি A, B দুটো matrix হয়, যাদের গুণ করা যায়, তবে সর্বদা $r(AB)$ হবে $r(A)$ এবং $r(B)$ এই দুজনের চেয়েই ছোটো বা সমান।

THEOREM

If A, B are any two matrices such that AB is defined, then

$$r(AB) \leq \min\{r(A), r(B)\}.$$

প্রমাণটা চেয়েছে নীচের অংকে।

Example 10: Let A, B be two matrices over the field $\mathbb{R}$ of real numbers such that AB is defined. Prove that

$$\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}.$$

[5] **(2006.13a)**

SOLUTION: এখানে অংকটা real field-এর জন্য করতে বললেও, যুক্তিটা যেকোনো field-এর জন্যই খাটে। প্রথমে এক কথায় প্রমাণের কায়দাটা বলে নিই। আমরা আগেই বলেছিলাম যে $\mathscr{C}(AB) \subseteq \mathscr{C}(A)$. সুতরাং dimension নিলেই পাচ্ছি $r(AB) \leq r(A)$. আবার এও দেখেছি যে $\mathscr{R}(AB) \subseteq \mathscr{R}(B)$. তা থেকে আসছে $r(AB) \leq r(B)$. ব্যস্‌!

> Let A be $m \times n$ and B be $n \times p$.
> Let the j-th column of A be $\vec{a}_j$.
> Then by definition of matrix multiplication
>
> $$AB = A[\vec{b}_1 \cdots \vec{b}_p] = [A\vec{b}_1 \cdots A\vec{b}_p].$$
>
> Thus the j-th column of AB is
>
> $$A\vec{b}_j = [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = b_{1j}\vec{a}_1 + \cdots + b_{nj}\vec{a}_n.$$
>
> Thus each column of AB is a linear combination of columns of A. Hence
>
> $$\mathscr{C}(AB) \subseteq \mathscr{C}(A).$$
>
> Thus
>
> $$r(AB) = \dim(\mathscr{C}(AB)) \leq \dim(\mathscr{C}(A)) = r(A).$$

এবার অন্যদিকটা দেখাব transpose নিয়ে--

> Again, $r(AB) = r((AB)') = r(B'A') \leq r(B')$ by the above argument.
> Since $r(B') = r(B)$, hence $r(AB) \leq r(B)$.
> Thus, $r(AB) \leq \min\{r(A), r(B)\}$, as required.

■

18.4.1 গুণ করলেও কখন rank কমে না?

একটা matrix-কে আরেকটা matrix দিয়ে গুণ করলে তার rank যে বেড়ে যেতে পারে না, সেটা আমরা দেখলাম। এবার কিছু বিশেষ ক্ষেত্র দেখব যেখানে গুণের পরেও rank-টা অপরিবর্তিত থাকে-- যদি Left থেকে full coLumn rank matrix দিয়ে, বা Right থেকে full Row rank matrix দিয়ে গুণ করা হয়। যেহেতু একটা nonsingular matrix একাধারে full column rank এবং full row rank, তাই কোনো nonsingular matrix দিয়ে যেকোনো দিকে থেকে গুণ করলেই rank অপরিবর্তিত থাকে।

> THEOREM
>
> Let $A_{m\times n}$ be any matrix. Then
>
> - for any full column rank $P_{p\times m}$, we have $r(PA) = r(A)$.
> - for any full row rank $Q_{n\times q}$, we have $r(AQ) = r(A)$.
> - hence, for any nonsingular matrices $R_{m\times m}$ and $S_{n]timesn}$, we have $r(RA) = r(AS) = r(A)$.

প্রমাণটা খুব একটা জটিল নয়। ধরো প্রথম অংশের প্রমাণটা। একটা দিক তো আগে থেকেই জানি--$r(PA) \leq r(A)$, কারণ গুণের ফলে rank বেড়ে যেতে পারে না। আবার P যেহেতু full coLumn rank, তাই ওর Left inverse আছে, ধরো B. মানে $BP = I$. সুতরাং $r(A) = r(B(PA)) \leq r(PA)$, যেহেতু PA-কে কিছু দিয়ে গুণ করলে rank বাড়তে পারে না। দুয়ে মিলে হল $r(PA) = r(A)$, ঠিক যেমনটা দরকার ছিল। একইভাবে অন্য অংশগুলোও হবে।
আরেকটা ক্ষেত্র আছে যেখানেও গুণ করলে rank কমে না। ধরো $A_{m\times n}$ যা খুশি একটা real matrix. এখানে matrix-টা real হওয়া খুবই দরকার (মানে field-টা $\mathbb{R}$ হওয়া চাই)। লক্ষ্য কর যে, সব সময়েই তুমি $A'A$ এবং AA' বার করতে পারবে, প্রথমটা হবে $n \times n$ আর দ্বিতীয়টা হবে $m \times m$ সাইজের। যদিও এখানে গুণ করা হয়েছে, তাও rank কমবে না, মানে $A, A'A$ এবং AA' এই তিনটে matrix-এর rank সমান হবে--

> THEOREM
>
> Let $A_{m\times n}$ be any matrix over field $\mathbb{R}$. Then
>
> $$r(A) = r(A'A) = r(AA').$$

Proof: আমরা প্রথমে $r(A) = r(A'A)$ দেখাব। বাকিটুকু পরে transpose নিলেই পেয়ে যাব। গুণ করলে rank বাড়ে না, তাই একটা দিক এমনিতেই পেয়ে যাচ্ছি--

> Since for any matrix B we have $r(AB) \leq r(A)$,
>
> $$\therefore r(A'A) \leq r(A). \qquad (*)$$

এবার অন্য দিকটা। এখানে একটা কৌশল লাগবে, যেটা চট্ করে ভেবে ওঠা মুশকিল। আমরা rank-এর বদলে প্রথমে nullity নিয়ে কাজ করব--

> Also we have $\mathscr{N}(A'A) \subseteq \mathscr{N}(A)$.
>
> ⟦ <u>Because:</u>
>
> Let $\vec{x} \in \mathscr{N}(A'A)$. Then $A'A\vec{x} = \vec{0}$.
> So $\vec{x}'A'A\vec{x} = \vec{x}'\vec{0} = 0$,
> or, $\vec{y}'\ y = 0$, where $\vec{y} = A\vec{x}$.
> Thus $y_1^2 + \cdots + y_m^2 = 0$.
> Since $\vec{y} \in \mathbb{R}^m$, $\therefore y_1 = \cdots = y_m = 0$.
> So $A\vec{x} = \vec{0}$, or $\vec{x} \in \mathscr{N}(A)$.
> Since $\vec{x} \in \mathscr{N}(A'A)$ was arbitrary, this proves
> $\mathscr{N}(A'A) \subseteq \mathscr{N}(A)$.
>
> ⟧
>
> Thus $\text{nullity}(A'A) \leq \text{nullity}(A)$.
>
> Now both $A'A$ and A have n columns. So
>
> $$r(A'A) = n - \text{nullity}(A'A) \geq n - \text{nullity}(A) = r(A).$$

সুতরাং দুয়ে মিলে--

> Combined with (*), this gives $r(A'A) = r(A)$.

প্রমাণের বাকি অংশটা খালি transpose লাগানোর অপেক্ষা--

> Applying this result to A' instead of A, we have
>
> $$r\big((A')'A'\big) = r(A'),$$
>
> or
>
> $$r(AA') = r(A).$$
>
> Hence
>
> $$r(A) = r(A'A) = r(AA'),$$
>
> as required.

[Q.E.D]

এই প্রমাণটা বুঝে থাকলে নীচের অংকটাও অনায়াসে হবে। এখানে কাজ হবে complex field-এর উপর, আর transpose-এর বদলে আসবে conjugate transpose, মানে A'-র জায়গায় A^*.

Exercise 26: For any matrix A over the field $\mathbb{C}$, show that

$$r(A) = r(A^*A) = r(AA^*).$$

■

DAY 19
আবার RREF

গত অধ্যায়ে আমরা reduced row echelon form (RREF)-এর কথা শিখেছিলাম। এবার এর আরেকটা প্রয়োগের কথা শিখব।

19.1 Inverse বার করা

আমরা শিখেছি যে, কোনো $A_{n\times n}$-কে nonsingular বলে যদি তার inverse থাকে। গতকাল আরো দেখেছি যে, এরকম matrix-এর rank হয় n, এবং এরা একই সঙ্গে full column rank আর full row rank হয়। কথা হল, এরকম একটা matrix দেওয়া থাকলে তার inverse আছে মাত্র জেনেই তো বসে থাকলে হবে না, সেটা তো বার করতে হবে। কী করে করবে? যদি matrix-টা 2×2 হয়, তবে একটা সহজ ফর্মুলা প্রথম অধ্যায়েই দেখেছিলাম--

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

সেখানে এটাও জেনেছিলাম যে, matrix-টা nonsingular হবে if and only if $ad-bc\neq 0$ হয়। সুতরাং ওই $\frac{1}{ad-bc}$-টাকে নিয়ে কোনো দুশ্চিন্তা নেই। কিন্তু যদি তোমাকে 10×10 সাইজের একটা nonsingular matrix দিয়ে তার inverse বার করতে বলি, তবে তো এরকম কোনো সহজ ফর্মুলা নেই! কিন্তু বাস্তবে এরকম কাজ প্রায়শঃই করতে হয়। তবে উপায়? উপায় সেই RREF এবং Gauss-Jordan elimination. কী করে বলি। তবে বোঝানোর সুবিধার জন্য 10×10 না নিয়ে 3×3 সাইজের matrix নিয়ে কাজ করব। একই কায়দা যেকোনো সাইজের nonsingular matrix-এর জন্যই কাজ করবে।

Example 11: Find the inverse of the matrix

$$A = \begin{bmatrix} -2 & 3 & 2 \\ 6 & 0 & 3 \\ 4 & 1 & -1 \end{bmatrix}$$

using RREF.

SOLUTION: RREF দিয়ে inverse বার করার প্রথম ধাপ হল matrix-টার পাশে একটা একই সাইজের identity matrix লাগিয়ে নেওয়া।

We start with

$$[A \quad I] = \left[\begin{array}{rrr|rrr} -2 & 3 & 2 & 1 & 0 & 0 \\ 6 & 0 & 3 & 0 & 1 & 0 \\ 4 & 1 & -1 & 0 & 0 & 1 \end{array}\right]$$

এখানে একটা খাড়া লাইন টেনে মূল matrix-টাকে identity অংশটার থেকে আলাদা রেখেছি। সেটা খালি বোঝার সুবিধার জন্য। তুমি মনে মনে পুরোটাকে একটা 3×6 সাইজের matrix বলেই ভেবো। এবার তোমার কাজ হবে Gauss-Jordan elimination লাগিয়ে এই matrix-টার RREF বার করা। মাঝের ধাপগুলো আর দেখাচ্ছি না। শেষ পর্যন্ত পাবে--

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{24} & \frac{5}{72} & \frac{1}{8} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{12} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{12} & \frac{7}{36} & -\frac{1}{4} \end{array}\right]$$

লক্ষ্য করো, শুরুতে খাড়া লাইনটার ডানদিকে identity ছিল, আর এখন identity রয়েছে লাইনটার বাঁদিকে। এতেই বোঝা যাচ্ছে যে, A ছিল nonsingular. যদি RREF-এ লাইনটার বাঁদিকে অন্যকিছু থাকত, তবে অমনি জোর দিয়ে বলতে পারতাম যে, A-টা nonsingular হতেই পারে না। যাই হোক, এখানে বাঁদিকে identity-ই আছে। তাই এবার লাইনটার ডানদিকে কী আছে দেখব। যেটা আছে, সেটাই হল A-র inverse, অর্থাৎ--

So the inverse is

$$A^{-1} = \begin{bmatrix} -\frac{1}{24} & \frac{5}{72} & \frac{1}{8} \\ \frac{1}{4} & -\frac{1}{12} & \frac{1}{4} \\ \frac{1}{12} & \frac{7}{36} & -\frac{1}{4} \end{bmatrix}.$$

■

লোকে যখন কম্পিউটার দিয়ে inverse বার করে, তখন কম্পিউটার সাধারণতঃ এই কায়দাটাই ব্যবহার করে।
কেন এই কায়দাটা কাজ করে সেটা এবার বোঝাব। কিন্তু তার আগে নীচের অংকে তুমি এই কায়দায় একটা matrix-এর inverse বার করার চেষ্টা করতে পারো।

Exercise 27: Invert the matrix

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 3 & 3 & 0 \\ 6 & 2 & 3 \end{bmatrix}$$

■

19.2 কেন কাজ করে?

গতকাল আমরা শিখেছি যে, একটা nonsingular matrix দিয়ে গুণ করলে rank বদলায় না। এই ব্যাপারটাই আসলে Gauss-Jordan elimination-এর প্রতি ধাপে ব্যবহৃত হচ্ছে।
Gauss-Jordan elimination করার সময়ে আমরা তিন রকম basic operation করি--

1. দুটো row-এর স্থানবিনিময় করা, বা swap করা,
2. একটা row-এর কোনো multiple নিয়ে অন্য কোনো row-এর সঙ্গে যোগ করা,
3. একটা row-কে কোনো nonzero সংখ্যা (মানে scalar) দিয়ে গুণ করা।

আমরা দ্বিতীয় অধ্যায়েই শিখেছিলাম যে, এই তিনটে কাজের ফলে matrix-টা যতই পরিবর্তিত হোক, তার row space কোনোমতেই বদলায় না (যদিও column space বদলাতেই পারে)। সুতরাং rank-ও অপরিবর্তিত থাকবে।
আমরা আগেই দেখেছি যে, nonsingular matrix দিয়ে গুণ করলে rank কখনও বদলায় না। আবার এখন দেখছি যে, basic operation করলেও rank বদলাচ্ছে না। ব্যাপারটা ঠিক কাকতালীয় নয়। Basic operation-রা আসলে বাঁদিক থেকে বিশেষ কিছু nonsingular matrix দিয়ে গুণ করা ছাড়া আর কিছুই নয়।

Example 12: যেকোনো একটা matrix নাও, যার দুটো row আছে (যতগুলো খুশি column থাকতে পারে)। এবার সেটাকে বাঁদিক থেকে এই matrix-টা দিয়ে গুণ করে দাও--

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

লক্ষ করবে যে, এর ফলে তোমার matrix-টার আর কিছুই পরিবর্তন হয় নি, খালি row দুটো স্থান বিনিময় করেছে!
SOLUTION: ধরো $A = \begin{bmatrix} 3 & 4 & 5 \\ 30 & 40 & 50 \end{bmatrix}$ নিলাম। তাহলে $SA = \begin{bmatrix} 30 & 40 & 50 \\ 3 & 4 & 5 \end{bmatrix}$ হবে। ■

সুতরাং বুঝতেই পারছ যে, S হল এমন একটা matrix, যা দিয়ে যেকোনো দুই row-ওয়ালা matrix-এর row-দুটোর স্থানবিনিময় করা যায়। যদি তার বদলে এমন একটা S চাইতাম, যা দিয়ে যেকোনো তিন row-ওয়ালা matrix-এর প্রথম

আর তৃতীয় row-এর স্থানবিনিময় করা যায়, তবে কী করে বার করতে? তার একটা সহজ কায়দা আছে। যেহেতু যেকোনো তিন row-ওয়ালা matrix-দের উপরেই ওর একই প্রভাব, সুতরাং একটা বিশেষ matrix নিয়ে কাজ করব--I_3, মানে 3×3 সাইজের identity matrix. এর প্রথম আর তৃতীয় row-র স্থানবিনিময় করলে কী হবে, তা তো জানিই। তার মানে আমরা চাইছি যেন

$$SI_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

হয়। কিন্তু identity matrix দিয়ে গুণ করলে তো আর কোনো matrix বদলায় না, তার মানে

$$S = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

এই S দিয়ে একটা তিন row-ওয়ালা matrix-কে গুণ করলেই প্রথম আর তৃতীয় row স্থানবিনিময় করে কিনা, সেটা নীচের অংকে নিজেই করে দেখো।

Exercise 28: দেখাও যে

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3n} \end{bmatrix} = \begin{bmatrix} a_{31} & \cdots & a_{3n} \\ a_{21} & \cdots & a_{2n} \\ a_{11} & \cdots & a_{1n} \end{bmatrix}.$$

■

একই কায়দা কাজ করবে যেকোনো basic operation-এর ক্ষেত্রেই। যা operation করতে চাও, identity matrix-এর উপরে করো। ফলে যা পাবে সেটাই হল সেই operation করার matrix.

Example 13: এমন একটা matrix S বার কর, যা দিয়ে বাঁদিক দিয়ে গুণ করলে যেকোনো তিন row-ওয়ালা matrix-এর প্রথম row-এর 4 গুণ দ্বিতীয় row-র সঙ্গে যোগ হয়ে যায়। যেমন আমরা চাই যেন

$$S\begin{bmatrix} 1 & 2 \\ 100 & 300 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 104 & 308 \\ 4 & 5 \end{bmatrix} \qquad (*)$$

হয়।

SOLUTION: প্রথমে I_3 নিয়ে শুরু করি--

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

এবার প্রথম row-এর 4-গুণ যোগ করি দ্বিতীয় row-এর সঙ্গে--

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

ব্যস্, এটাই আমাদের প্রয়োজনীয় S. এবার (*)-এর গুণটা করেই দ্যাখো। ঠিক মিলবে! ■

Example 14: চারটে row-ওয়ালা matrix-দের row-এর উপরে একটা basic operation হল দ্বিতীয় row-টাকে 5

গুণ করা। এর জন্য matrix-টা কী হবে?
SOLUTION: I_4-এর উপরে এই operation-টা লাগালে পাই--

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

এটাই আমাদের প্রয়োজনীয় matrix. ■

আমরা এতক্ষণ খালি row-গুলোর উপরে basic operation করছিলাম, একইভাবে column-দের উপরেও করা যায়। তার জন্য nonsingular matrix দিয়ে গুণটা করতে হবে ডানদিক থেকে। Matrix-গুলো এখানেও পাওয়া যাবে identity matrix-এর উপর basic operation-টা লাগিয়ে। আরেকটা পথ হল প্রথমে column-গুলোকে row-তে পরিণত করে নেওয়া। একটা উদাহরণ দেখি--

Example 15: এমন একটা matrix বার কর, যা দিয়ে তিনটে column-ওয়ালা কোনো matrix-কে গুণ করলে প্রথম column-এর 5-গুণ তৃতীয় column-এ যোগ হয়ে যায়।
SOLUTION: ধরো যা খুশি একটা তিন-column-ওয়ালা matrix নিলাম A. আমরা row-গুলোর উপর basic operation-এর জন্য matrix বার করা জানি। তাই প্রথমে A' বার করে নিই, তাহলে column-গুলো row হয়ে যাবে। এবার A'-এর প্রথম row-এর 5-গুণ যোগ করতে হবে তৃতীয় row-এর সাথে। এই কাজের জন্য matrix-টা পাব I_3 থেকে--

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{new\ \ R_3 = R_3 + 5R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix} = S.$$

সুতরাং row-এর উপর basic operation-টা করার পরে পেলাম SA'. এবার transpose করে দিলেই column-রা ফের column-এর জায়গায় ফিরে যাবে। সব মিলিয়ে হল

$$(SA')' = AS'.$$

তার মানে আমাদের প্রয়োজনীয় matrix-টা হল S', এবং গুণটা করতে হবে ডানদিক থেকে। ■

এই উদাহরণটা খুবই শিক্ষণীয়। Column-এর উপর যেকোনো basic operation-এর ক্ষেত্রেই একই যুক্তি খাটবে। Row-এর জন্য যে matrix দিয়ে বাঁদিক দিয়ে গুণ করতে হয়, তার transpose দিয়ে ডানদিক থেকে গুণ করলেই সেই একই basic operation হবে column-এর উপর।
Row/column-এর উপর basic operation করার জন্য matrix-দের বলে **elementary matrix**. এরা খুবই কাজের জিনিস হলেও তিনটে basic operation-এর জন্য elementary matrix-দের আলাদা করে কোনো প্রচলিত নাম নেই। একেকটা বই একেক রকম নাম ব্যবহার করে। আমরা এই নাম কয়টা ব্যবহার করব--

1. i-th আর j-th row-এর স্থানবিনিময় করার matrix-কে বলব $S(i,j)$. এখানে $i \neq j$.
2. i-th row-এর α-গুণ নিয়ে j-th row-এর সাথে যোগ করে দেওয়ার matrix-কে বলব $A(i,\alpha,j)$. এখানে $i \neq j$, আর α হল যে কোনো scalar.
3. i-th row-কে α দিয়ে গুণ করার matrix-এর নাম দেব $M(i,\alpha)$. এখানে α যেকোনো nonzero scalar.

একটু চিন্তা করলেই দেখবে যে এই তিন ধরণের elementary matrix-ই nonsingular. শুধু তাই নয়, এদের inverse-রাও আবার elementary matrix-ই হবে। সেটাই এবার আলোচনা করব।

Example 16: ধরো তুমি একটা matrix-এর i-th এবং j-th row-এর স্থানবিনিময় করেছ। এমন সময়ে খবর এল যে, সেটা করার দরকার নেই, তবে চট্ করে গোড়ার matrix-টায় কী করে ফিরে যাবে?
SOLUTION: i-th আর j-th row দুটোকে আরেকবার স্থানবিনিময় করে দিলেই ওরা দুজনেই স্বস্থানে ফিরে যাবে। তার মানে বাঁদিক থেকে পরপর দুইবার $S(i,j)$ দিয়ে গুণ করলে ওরা পরস্পর কাটাকাটি হয়ে যায়। অর্থাৎ $S(i,j)$-এর inverse হল $S(i,j)$ নিজেই! ■

Example 17: একটা matrix-এর i-th row-কে $\alpha \neq 0$ দিয়ে গুণ করে ফেলেছো। এবার সেখান থেকে মূল matrix-টায় ফিরে যাওয়ার উপায় কী?
SOLUTION: ওই row-টাকে α দিয়ে ভাগ করে দিলেই চলবে, মানে $\frac{1}{\alpha}$ দিয়ে গুণ করে দিতে হবে। সুতরাং $M(i,\alpha)$-এর inverse হল $M\left(i,\frac{1}{\alpha}\right)$. ■

Exercise 29: একটা matrix-এর i-th row-এর α গুণ তুমি যোগ করে দিয়েছ j-th row-এর সঙ্গে। এখান থেকে মূল matrix-টায় ফিরে যাবার উপায়? এ থেকে বলো $(A(i,\alpha,j))^{-1}$ কী হবে। ■

এবার বুঝতে পারবে কেন Gauss-Jordan elimination দিয়ে একটা nonsingular matrix-এর inverse বার করা যাচ্ছিল। প্রথম কথা, যেকোনো nonsingular matrix-এর RREF-ই হল identity matrix. কারণ, যদি $A_{n\times n}$ হয় nonsingular, তবে $\mathscr{R}(A) = F^n$. এদিকে $\mathscr{R}(I_n)$-ও হল F^n. দেখাই যাচ্ছে যে, I_n অবশ্যই RREF চেহারায় আছে। আর আমরা জানি যে, যেকোনো matrix-এরই RREF-টা unique হয়। তার মানে A-র RREF ওই I_n-ই হতে বাধ্য। সুতরাং এটুকু অন্ততঃ বোঝা গেল যে, খাড়া লাইনটার বাঁদিকে কেন identity আসতে বাধ্য। এবার বুঝব, ডানদিকে কেন A^{-1} আসছে।
আমরা প্রতিটা ধাপে যখন কিছু basic operation করছি, তখন বাঁদিক থেকে কিছু nonsingular matrix দিয়ে গুণ করছি। আমরা প্রথম অধ্যায়ে যখন partitioned matrix-এর গুণ শিখেছিলাম, তখন শিখেছিলাম যে--

$$P\left[\begin{array}{c|c} Q & R \end{array}\right] = \left[\begin{array}{c|c} PQ & PR \end{array}\right].$$

সেই যুক্তি অনুযায়ী আমরা প্রথম ধাপে যদি বাঁদিক থেকে E_1 দিয়ে গুণ করি, তবে প্রথম ধাপের শেষে পাব--

$$E_1\left[\begin{array}{c|c} A & I \end{array}\right] = \left[\begin{array}{c|c} E_1A & E_1 \end{array}\right].$$

এরপর দ্বিতীয় ধাপে যদি E_2 দিয়ে গুণ করি, তবে সেই ধাপের শেষে পাব--

$$E_2\left[\begin{array}{c|c} E_1A & E_1 \end{array}\right] = \left[\begin{array}{c|c} E_2E_1A & E_2E_1 \end{array}\right].$$

এইভাবে চলতে চলতে একেবারে শেষে পাব $\left[\begin{array}{cc} EA & E \end{array}\right]$, যেখানে $E = E_nE_{n-1}\cdots E_2E_1$. একেবারে শেষে খাড়া লাইনটার বাঁদিকে যে I থাকবে, সে তো দেখেইছি। তার মানে $EA = I$. সুতরাং $E = A^{-1}$. অর্থাৎ খাড়া লাইনের ডানদিকের matrix-টাই হচ্ছে A^{-1}.
একটা উদাহরণ নিয়ে দেখা যাক।

Example 18: ধরো A নিলাম এইরকম--

$$A = \begin{bmatrix} -2 & 3 & 2 \\ 6 & 0 & 3 \\ 4 & 1 & -1 \end{bmatrix}.$$

তবে $\begin{bmatrix} A & I \end{bmatrix}$ থেকে শুরু করে দেখাও যে, প্রতি ধাপেই আমরা $\begin{bmatrix} EA & E \end{bmatrix}$ জাতীয় কিছু একটা পাচ্ছি, মানে খাড়া লাইনের ডান দিকের matrix-টা দিয়ে A-কে গুণ করলে বাঁদিকের matrix-টা পাওয়া যায়।

SOLUTION: প্রথম ধাপের শেষে--

$$\left[\begin{array}{ccc|ccc} 1 & -\frac{3}{2} & -1 & -\frac{1}{2} & 0 & 0 \\ 0 & 9 & 9 & 3 & 1 & 0 \\ 0 & 7 & 3 & 2 & 0 & 1 \end{array}\right]$$

সত্যই এখানে

$$\begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 3 & 2 \\ 6 & 0 & 3 \\ 4 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{3}{2} & -1 \\ 0 & 9 & 9 \\ 0 & 7 & 3 \end{bmatrix}.$$

দ্বিতীয় ধাপের শেষে--

$$\left[\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{2} & 0 & \frac{1}{6} & 0 \\ 0 & 1 & 1 & \frac{1}{3} & \frac{1}{9} & 0 \\ 0 & 0 & -4 & -\frac{1}{3} & -\frac{7}{9} & 1 \end{array}\right]$$

এখানেও একইভাবে গুণ করে দেখো, ঠিক মিলে যাবে।

তৃতীয় ধাপের শেষে আসে--

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{24} & \frac{5}{72} & \frac{1}{8} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{12} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{12} & \frac{7}{36} & -\frac{1}{4} \end{array}\right]$$

এখানেও মিলে যেতে বাধ্য। ■

DAY 20
System of linear equations (part 1)

20.1 Theory

আমরা স্কুলে থাকতেই এরকম equation-এর কথা শিখি--

$$\begin{aligned} 2x + 3y &= 7 \\ x - y &= 1. \end{aligned}$$

এটা হল যাকে বলে একটা **system of linear equations**. এখানে "system" বলতে "একটা গুচ্ছ" বোঝাচ্ছি। সেই গুচ্ছের প্রতিটা সদস্য হল একেকটা linear equation, মানে $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ জাতীয় জিনিস, যেখানে $x_1, ..., x_n$ হল unknown, আর $a_1, ..., a_n$ হল কিছু সংখ্যা, আর b-ও একটা সংখ্যা। এখানে "linear" বলার মানে x_i-দের কোনো square, cube, sin, cos ইত্যাদি নেওয়া চলবে না, বা ওদের নিজেদের মধ্যে গুণভাগ করাও যাবে না। সুতরাং $xy = 4$ মোটেই একটা linear equation নয়, কিন্তু $x - 4y + 4z = 4$ অবশ্যই একটা linear equation.

কিছু linear equation-এর একটা system-কে আমরা চাইলে matrix দিয়ে লিখতে পারি। যেমন, শুরুতে যে system-টা দিয়েছিলাম, সেটাকে matrix দিয়ে লিখলে হয়--

$$\begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}.$$

এখানে $\begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}$-কে বলব **coefficient matrix**. এইভাবে লিখলে যেকোনো system of linear equations-এরই general রূপ হয়--

$$A_{m\times n}\vec{x}_{n\times 1} = \vec{b}_{m\times 1}.$$

এখানে m হল equation-এর সংখ্যা, আর n হল unknown-এর সংখ্যা। যেমন system-টা

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 10 \\ 20 \end{bmatrix}}_{\vec{b}}$$

হলে equation আছে দুটো, আর unknown আছে চারটে।
আজকের ক্লাস আমরা শুরু করেছিলাম এই system-টা দিয়ে--

$$\begin{aligned} 2x + 3y &= 7 \\ x - y &= 1. \end{aligned}$$

ছোটোবেলায় যেভাবে সমাধান করতাম, সেভাবে করলে উত্তর পাবে $x = 2, y = 1$. এটাই একমাত্র উত্তর, অর্থাৎ unique solution.
সব সময়ে অবশ্য কাজটা এত সহজে হয় না। যেমন ধরো এই system-টা--

$$\begin{aligned} x + 2y &= 3 \\ 2x + 4y &= 5. \end{aligned}$$

একটু খুঁটিয়ে দেখলেই দেখবে যে এর কোনো সমাধানই সম্ভব নয়, কারণ equation দুটো কিছুতেই একই সঙ্গে ঠিক হতে পারে না! কারণ--

> প্রথম equation-টা বলছে $x + 2y = 3$. তাই যদি হয় তবে $2x + 4y = 2(x + 2y) = 2 \times 3 = 6$ হতে বাধ্য। অথচ দ্বিতীয় equation-টা দাবী করছে যে $2x + 4y$ নাকি 5.

এই রকম system-কে বলে **inconsistent** (বা সামঞ্জস্যবিহীন)।
এবার আরেক রকম উদাহরণ দ্যাখো--

$$\begin{aligned} x + 2y &= 3 \\ 2x + 4y &= 6. \end{aligned}$$

এটা বস্তুতঃ আগের system-টাই, খালি শেষের 5-টাকে ঠিক করে এবার 6 করে দিয়েছি। সুতরাং এই system-টা আর inconsistent নেই। এবার কি সমাধান করে x, y বার করতে পারবে? ভালো করে দেখলেই বুঝবে যে এখানে আসলে একটাই equation. অন্যটা ওটারই একটা multiple মাত্র! একবার যেই প্রথম equation-টা বলে দিল যে $x + 2y = 3$, তখন বলাই বাহুল্য যে $2x + 4y = 6$ হবে! কিন্তু একটা মাত্র equation থেকে x, y দুজনকে কী করে বার করব? করা যাবে না। যেটা করা যাবে সেটা হল x আর y-এর মধ্যে একজনকে অন্যজনের মাধ্যমে প্রকাশ করা। যেমন, y-কে x দিয়ে প্রকাশ করলে হয়--

$$y = \tfrac{1}{2}(3 - x).$$

ব্যস্, এর বেশী আর কিছুই বলা যায় না। তার মানে x তুমি যা খুশি নাও, খালি y-কে সেই অনুযায়ী নিলেই চলবে। এইভাবে তুমি নানারকম solution পেতে পারো, যেমন $x = 1$ বসালে পাবে $y = 1$, সুতরাং $(x, y) = (1, 1)$ হল একটা solution. আবার $x = 0$ বসালে পাবে $y = \frac{3}{2}$, অতএব $(x, y) = \left(0, \frac{3}{2}\right)$ হবে আরেকটা solution, ইত্যাদি। এরকম যত রাজ্যের solution-এর একটা লম্বা তালিকা না বানিয়ে আমরা এক কথায় লিখে দিই--

The system has general solution $\left(x, \frac{1}{2}(3-x)\right)$, where x is arbitrary.

আমরা তিনরকমের সম্ভাবনার কথা শিখলাম--

1. unique solution
2. inconsistent
3. more than one solution.

বুঝতেই পারছ এর বাইরে আর সম্ভাবনা থাকতে পারে না, solution-এর সংখ্যা হয় 0, নয়তো 1, আর নয়তো > 1. সবার শেষে যে উদাহরণটা করলাম, সেখানে একাধিক solution ছিল। শুধু একাধিক নয়, অসংখ্য solution ছিল, কারণ x-কে যা খুশি (মানে যে কোনো real number) নেওয়া যাচ্ছিল। এখানে real number-এর প্রসঙ্গটা এল কেন? কারণ আমরা ধরে নিয়েছিলাম যে, solution বলতে আমরা $x, y \in \mathbb{R}$ খুঁজছি। স্কুলে পড়ার সময়ে অবশ্য বইতে আলাদা করে $\mathbb{R}$-এর উল্লেখ করা থাকত না, কিন্তু সেটাই বুঝে নেওয়া হত। Linear algebra-র ভাষায় বলতে হলে বলব যে, এখানে আমরা যে field-এর উপর কাজ করছি সেটা হল $\mathbb{R}$. যদি $\mathbb{R}$-এর বদলে $\mathbb{C}$-এর উপর কাজ করতাম তবে general solution-টা হত $\left(x, \frac{1}{2}(3-x)\right)$, যেখানে $x \in \mathbb{C}$. এখানেও কিন্তু solution-এর সংখ্যা infinite. সুতরাং solution-এর সংখ্যা যে infinite, সেটা আসছে field-টা infinite বলে। প্রথম অধ্যায়ে শিখেছিলাম যে, সব field-ই infinite নয়, অনেক finite field-ও আছে। যদি সেরকম কোনো finite field নিতাম, তবে কিন্তু একাধিক solution হলেও অসংখ্য solution পেতাম না।

আচ্ছা, মনে করো আমরা একটা infinite field-এর উপরে কাজ করছি, তবে কি এমন একটা system সম্ভব, যার একাধিক solution আছে, কিন্তু অসংখ্য solution নেই? যেমন মনে করো ঠিক দুটোই solution আছে? এই প্রশ্নটার উত্তর কিন্তু স্কুলে যেটুকু শিখেছিলে, তার ভিত্তিতে দেওয়া সহজ নয়। এখানে linear algebra-র শরণাপন্ন হতে হবে, এবং system-টাকে matrix দিয়ে লিখলে সুবিধা হবে। প্রশ্নটা শুনতে সহজ লাগলেও এর উত্তর খুঁজতে গিয়ে আমরা এমন একটা theorem-এর দেখা পাব যেটা ভীষণই গুরুত্বপূর্ণ। ধরো system-টা হল

$$A\vec{x} = \vec{b},$$

যার একাধিক solution (মানে অন্ততঃ দুটো solution) আছে। ধরো দুটো solution-এর নাম দিলাম $\vec{x}_0$ আর $\vec{x}_1$. তার মানে $A\vec{x}_0 = \vec{b}$ এবং $A\vec{x}_1 = \vec{b}$. যদি একটা থেকে আরেকটাকে বিয়োগ করে দিই, তবে পাব

$$A(\vec{x}_0 - \vec{x}_1) = \vec{b} - \vec{b} = \vec{0}.$$

সুতরাং $\vec{x}_0 - \vec{x}_1 \in \mathscr{N}(A)$. তার মানে যেকোনো দুটো solution-এর মধ্যে পার্থক্যটা সব সময়েই $\mathscr{N}(A)$-র মধ্যে থাকতে বাধ্য। এর উল্টো দিকটাও কিন্তু ঠিক--যদি $\vec{x}_0$ যেকোনো solution হয়, আর $\vec{v} \in \mathscr{N}(A)$ হয় তবে $\vec{x}_0 + \vec{v}$-ও একটা solution হবেই। কারণ--

$A\vec{x}_0 = \vec{b}$ আর $A\vec{v} = \vec{0}$ দেওয়া আছে। তা থেকে পাচ্ছি $A(\vec{x}_0 + \vec{v}) = A\vec{x}_0 + A\vec{v} = \vec{b} + \vec{0} = \vec{b}.$

এই দুই দিক মিলিয়ে আমরা একটা খুব গুরুত্বপূর্ণ জিনিস পেলাম--

THEOREM

Let $A\vec{x} = \vec{b}$ have one solution $\vec{x}_0$. Then the set of all solutions is

$$\vec{x}_0 + \mathscr{N}(A).$$

Exercise 30: Find the dimension and basis of the solution space of the system of equations:

$$\begin{aligned} x-2y+z &= 0\\ x-2y-z &= 0\\ 2x-4y-5z &= 0. \end{aligned}$$

[5] **(2011.12a)**

HINT:

এই অংকে solution space মানে যাবতীয় solution-এর set. এখানে $\vec{b}=\vec{0}$. তাই solution space-টা এখানে $\mathscr{N}(A)$ ছাড়া কিছুই নয়। অবশ্য এই অংকটার জন্য অত কিছু না ভাবলেও চলবে, সরাসরি করে গেলেই হবে। ■

মোট কতগুলো solution আছে সেটা সম্পূর্ণই নির্ভর করছে $\mathscr{N}(A)$-র সাইজের উপর। আমরা জানি যে $\mathscr{N}(A)$ হল একটা subspace. যদি সেটা trivial subspace হয়, তবে তার সাইজ হল 1, সেক্ষেত্রে $\vec{x}_0$-টাই একমাত্র solution হবে। যদি $\mathscr{N}(A)$-টা trivial না হয়, তবে ওর সাইজ নির্ভর করবে যে field-এর উপর কাজ হচ্ছে, তার সাইজের উপর। যদি field-টা infinite হয়, তবে $\mathscr{N}(A)$-ও হবে infinite সাইজের, সুতরাং infinitely many solution থাকবে। যদি field-টা finite হয়, তবে $\mathscr{N}(A)$-র সাইজও finite হবে (কিন্তু >1) এবং সেক্ষেত্রে একাধিক solution থাকবে, কিন্তু finite-সংখ্যক।
সুতরাং আমাদের প্রশ্নের উত্তর পেয়ে গেলাম--কোনো infinite সাইজের field-এর উপর এমন system অসম্ভব, যার ঠিক দুটোই solution আছে।
এবার পরবর্তী প্রশ্ন--কী করে বুঝব একটা system আদৌ consistent কিনা? একটা পথ অবশ্যই solve করার চেষ্টা করা। যদি কোথাও আটকে যাই, তবে inconsistent, নইলে consistent. কিন্তু কাজটা তার চেয়ে সহজে করা যায় linear algebra ব্যবহার করে। এর জন্য আমাদের প্রথম উদাহরণটা দেখি--

$$\begin{bmatrix} 2 & 3\\ 1 & -1 \end{bmatrix}\begin{bmatrix} x\\ y \end{bmatrix}=\begin{bmatrix} 7\\ 1 \end{bmatrix}.$$

এখানে আমরা এমন $x,y\in\mathbb{R}$ খুঁজছি যাতে

$$\begin{aligned} 2\times x+3\times y &= 7\\ 1\times x-1\times y &= 1 \end{aligned}$$

হয়। যাবতীয় x-গুলোকে কমন নিয়ে, আর y-গুলোকে আলাদা করে কমন নিয়ে পাই--

$$x\begin{bmatrix} 2\\ 1 \end{bmatrix}+y\begin{bmatrix} 3\\ -1 \end{bmatrix}=\begin{bmatrix} 7\\ 1 \end{bmatrix}$$

হয়। এইটা ভালো করে বুঝে নাও। সুতরাং আমরা যেন জানতে চাইছি যে $(7,1)$-কে A-র column-গুলোর linear combination হিসেবে লেখা যায় কিনা। লেখা গেলে, কী করে লেখা যায়, তা থেকে $\vec{x}$ বেরোবে। এখানে সেটা নিয়ে আমাদের মাথাব্যথা নেই। আপাততঃ প্রশ্ন হল system-টা consistent কিনা, মানে আদৌ linear combination হিসেবে লেখা যায় কিনা, অর্থাৎ $(7,1)$-টা $\mathscr{C}(A)$-এর মধ্যে আছে কি না। এই সুতরাং সব মিলিয়ে পেলাম--

THEOREM

A system $A\vec{x}=\vec{b}$ is consistent if and only if $\vec{b}\in\mathscr{C}(A)$.

এইটা এবং আগের theorem-টার সৌন্দর্য্যটা লক্ষ কর। একটা system-কে স্কুলের মত করে না লিখে matrix দিয়ে লিখলে বোঝা যায় ওর মধ্যে দুটো অংশ আছে A আর $\vec{b}$. এই দুজন কীভাবে system-টার আচরণ নিয়ন্ত্রণ করছে, সেটা ফুটে

ওঠে এই দুটো theorem-এ। Consistent হবে নাকি inconsistent হবে, সেটা নির্ভর করে A এবং $\vec{b}$-এর মধ্যে একটা সম্পর্কের উপর, $\vec{b} \in \mathscr{C}(A)$ হচ্ছে কিনা। কিন্তু যদি একবার consistent হয়ে যায়, তবে কতগুলো solution থাকবে সেটা ঠিক করার দায়িত্ব পুরোটাই A-র, সেখানে $\vec{b}$-র কোনোই ভূমিকা নেই। এই ব্যাপারটা আরও একধাপ এগিয়ে নিয়ে গেলে আরেকটা সুন্দর theorem পাওয়া যায়। তার জন্য একটা কথা মনে করিয়ে দিই। যদি $\vec{v}_1, ..., \vec{v}_n$ কিছু vector হয়, আর বলা থাকে যে $\vec{u} \in \text{span}\{\vec{v}_1, ..., \vec{v}_n\}$, তার মানে $\vec{u}$-কে এইভাবে লেখা যাবে--

$$\vec{u} = a_1\vec{v}_1 + \cdots + a_v v_n,$$

যেখানে a_i-রা সবাই scalar. এমনিতে a_i-গুলোকে একাধিক ভাবে নেওয়া যেতে পারে, কিন্তু যদি $\{\vec{v}_1, ..., \vec{v}_n\}$ একটা linearly independent set হয়, তবে a_i-দের খালি একভাবেই নেওয়া সম্ভব। এবং উল্টো দিকটাও সত্যি--যদি দেখা যায় যে অন্ততঃ একটা $\vec{u} \in \text{span}\{\vec{v}_1, ..., \vec{v}_n\}$-এর জন্যও a_i-গুলো unique হয়ে যাচ্ছে, তবে অবশ্যই $\{\vec{v}_1, ..., \vec{v}_n\}$ একটা linearly independent set হতে বাধ্য। এই কথাটা আমরা আগেই শিখেছিলাম, এবার কাজে লাগাবার সময় এসেছে।

আমরা দেখেছি যে, $A\vec{x} = \vec{b}$ দেওয়া থাকলে সেটা সমাধান করা মানে $\vec{b}$-কে A-র column-গুলোর linear combination হিসেবে লেখা, কারণ ওই linear combination-এর coefficient-গুলো দিয়েই $\vec{x}$ তৈরী। যদি $\vec{b} \in \mathscr{C}(A)$ হয়, তবে কখন solution-টা unique হবে? ঠিক যখন A-র column-গুলো মিলে একটা linearly independent set তৈরী করবে। অতএব আমরা তৃতীয় একটা theorem পেলাম, যেটা বলছে যে, solution যদি থাকে, তবে তার uniqueness নির্ভর করে খালি A-র উপরে, $\vec{b}$-এর উপরে নয়।

THEOREM

A consistent system $A\vec{x} = \vec{b}$ has unique solution if and only if the columns of A form a linearly independent set, *i.e.*, if A is full column rank.

A-টা full column rank হলে কী হয় শেখা গেল। আন্দাজ করা যায় যে, full row rank-এর জন্যও কিছু একটা theorem থাকবে, মানে যদি A-র row-গুলো linearly independent হয়। Theorem-টা সোজাই--

THEOREM

Let A be a full row rank matrix. Then for any $\vec{b}$ the system

$$A\vec{x} = \vec{b}$$

must be consistent.

প্রমাণ অতি সহজ। ধরো A-র সাইজ হল $m \times n$. তাহলে এখানে $r(A) = m$. তার মানে $\mathscr{C}(A)$-এর dimension হল m. তাই $\mathscr{C}(A) = F^m$, কারণ--

> $\mathscr{C}(A) \subseteq F^m$ তো আছেই, এখন আবার $\dim(\mathscr{C}(A)) = \dim(F^m)$-ও পাওয়া যাচ্ছে। এই যুক্তিটা আমরা ইতিপূর্বেই দেখেছি।

অতএব তুমি যাই $\vec{b} \in F^m$ নাও না কেন, $\vec{b} \in \mathscr{C}(A)$ না হয়ে যায় না, তাই $A\vec{x} = \vec{b}$ অতি অবশ্যই consistent হবে। এই দুটো theorem-এর মধ্যে একটা মিল আছে--

- যদি A-টা full column rank হয়, তবে solution-এর সংখ্যা হবে ≤ 1. (কারণ যদি inconsistent হয়, তবে solution-এর সংখ্যা হবে 0, আর consistent হলে যেহেতু solution-টা unique হতে বাধ্য, তাই তখন solution-এর সংখ্যা হবে 1).

- যদি A-টা full row rank হত, তবে solution-এর সংখ্যা হবে ≥ 1 (যেহেতু consistent, তাই অন্ততঃ একটা solution থাকবেই)।

যদি A-টা একই সঙ্গে full column rank আর full row rank দুইই হয়? সেক্ষেত্রে অবশ্যই A-কে square হতে হবে এবং nonsingular হতে হবে। তখন solution-এর সংখ্যা একাধারে ≤ 1 এবং ≥ 1 হতে হবে, মানে ঠিক একটাই solution থাকবে। সেটা বলাই বাহুল্য কারণ A যদি nonsingular হয়, তবে $A\vec{x} = \vec{b}$-এর unique solution অবশ্যই $A^{-1}\vec{b}$.

20.2 System-এর চরিত্র বিশ্লেষণ

আমরা দেখলাম--যেকোনো system of linear equations-ই তিনটে ধরণের মধ্যে একটা হয়--

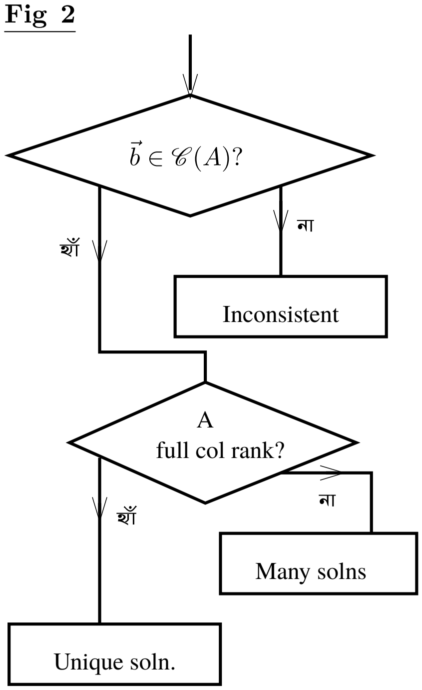
1. inconsistent, মানে কোনো solution নেই,
2. unique solution, মানে ঠিক একটাই solution আছে,
3. একাধিক solution আছে। ঠিক কতগুলো থাকবে, সেটা নির্ভর করবে field-টার উপর। যদি field-এর সাইজটা infinite হয়, তবে এক্ষেত্রে solution-ও থাকবে infinitely many. যদি field-টা finite হয়, তবে solution-ও থাকবে finite-সংখ্যক (কিন্তু একাধিক)।

ধরো তোমাকে একটা system of linear equations দেওয়া আছে--

$$A\vec{x} = \vec{b}.$$

তোমার কাজ হল পরীক্ষা করে বলা সেটা কোন ধরণের। আমরা যে theorem-গুলো শিখেছি তা দিয়েই কাজটা করা যাবে। ধাপগুলো দেখিয়েছি Fig 2-এ। প্রথমে দেখতে হবে $\vec{b} \in \mathscr{C}(A)$ হচ্ছে কিনা। যদি না হয়, তবে সেখানেই ল্যাঠা চুকে গেল, system-টা অবশ্যই inconsistent হবে। যদি $\vec{b} \in \mathscr{C}(A)$ হয়, তবে দেখতে হবে A-টা full column rank কিনা। যদি হয়, তবে unique solution, নইলে একাধিক solution.

Fig 2



তার মানে আমাদের দুটো জিনিস পরীক্ষা করে দেখতে হচ্ছে--,এক, $\vec{b} \in \mathscr{C}(A)$ হচ্ছে কিনা, আর দুই, A-টা full column rank কিনা। প্রশ্ন হল--এই দুটো পরীক্ষা করব কীভাবে? হয়তো আন্দাজ করতে পারছ যে, RREF কাজে লাগবে। হ্যাঁ, শুধু তাই নয়, দুটো পরীক্ষাই করা যাবে rank ব্যবহার করে। তার জন্য একটা নতুন জিনিস বানিয়ে নিলে সুবিধা হবে, যার নাম **augmented matrix** (অগ্‌মেন্টেড ম্যাট্রিক্স)।

20.2.1 Augmented matrix

আমাদের system-টা ছিল $A\vec{x} = \vec{b}$. যদি A-র পাশে $\vec{b}$-কে জুড়ে দিই, তবে পাব

$$\begin{bmatrix} A & \vec{b} \end{bmatrix}.$$

একেই বলে এই system-টার augmented matrix. ইংরাজীতে "augment" করা মানে হল "বাড়ানো"। যেহেতু A-র সঙ্গে $\vec{b}$ লাগিয়ে তাকে বাড়ানো হয়েছে, সেইজন্যই এইরকম নাম।

Example 19: যদি আমাদের system-টা হয়

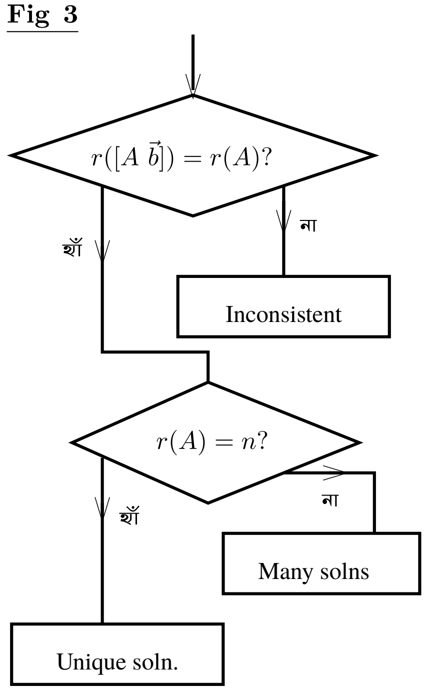
$$\begin{aligned} x + 2y &= 5 \\ 3x - y &= 1 \end{aligned}$$

তবে এর augmented matrix-টা কী হবে?

SOLUTION: এখানে $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$, আর $\vec{b} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$. এদেরকে পাশাপাশি জুড়ে দিলেই পাব এই augmented matrix-টা--

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & -1 & 1 \end{array}\right].$$

Fig 3



লক্ষ কর আমরা A আর $\vec{b}$-এর মাঝখানে একটা লাইন টেনে দিয়েছি। এটা না দিলেও মহাভারত অশুদ্ধ হত না, তবে এটাই দস্তুর, যাতে খেয়াল থাকে যে লাইনের বাঁদিকে A আর ডানদিকে $\vec{b}$. ■

এবার লক্ষ কর যে A হল $[A \quad \vec{b}]$-এর একটা submatrix. সুতরাং $r(A) \leq r([A \quad \vec{b}])$ হতে বাধ্য। যদি $\vec{b} \in \mathscr{C}(A)$ হয়, তার মানে হল $\vec{b}$ আসলে A-এর column-দের linear combination. সেক্ষেত্রে $r([A \quad \vec{b}]) = r(A)$ হবে। নইলে হবে $r(A) < r([A \quad \vec{b}])$. সুতরাং consistent কিনা তার পরীক্ষা হয়ে যাচ্ছে $r([A \quad \vec{b}])$ আর $r(A)$ সমান কিনা সেটা দেখলেই। এর পরের পরীক্ষা হল A-টা full column rank কিনা। ধরো A হল $m \times n$. তবে full column rank হওয়া মানে $r(A) = n$ হওয়া। সুতরাং Fig 2-এর কাজটাই rank ব্যবহার করে করলে পাব Fig 3.
অনেকটা theory শিখলাম। একটু হাতে কলমে প্রয়োগ করে দেখি।

Example 20: ধরো একটা system আছে $A_{m\times n}\vec{x} = \vec{b}$. নীচের প্রতিক্ষেত্রে বল system-টা কীরকম হবে, inconsistent হবে, নাকি unique solution থাকবে, নাকি একাধিক solution থাকবে।

1. $r(A) = 4$ এবং $r\left(\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]\right) = 5.$
1. $r(A) = n = 5$ এবং $r\left(\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]\right) = 5.$
1. $n = 10,\ r(A) = 5$ এবং $r\left(\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]\right) = 5.$

SOLUTION:

1. এখানে $r(B) < r\left(\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]\right)$. তাই $\vec{b} \notin \mathscr{C}(A)$. অতএব inconsistent.
2. এখানে $r(B) = r\left(\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]\right)$. তাই $\vec{b} \in \mathscr{C}(A)$. অতএব consistent. এদিকে $r(A) = n$, তাই full column rank. সুতরাং unique solution.
3. এখানেও consistent, কিন্তু full column rank নয় (যেহেতু $r(A) < n$)। তাই একাধিক solution.

■

Exercise 31: ধরো একটা system আছে $A_{m\times n}\vec{x} = \vec{b}$. নীচের প্রতিক্ষেত্রে বল system-টা কীরকম হবে, inconsistent হবে, নাকি unique solution থাকবে, নাকি একাধিক solution থাকবে।

1. $n = 6,\ r(A) = 6$ এবং $r\left(\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]\right) = 6.$
2. $m = 4,\ n = 5$ এবং $r(A) = 4$
3. $n = 10$ এবং $r\left(\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]\right) = 11.$

■

স্কুলে থাকতে আমরা অনেক সময়ে কিছু জিনিস ভেবে নিই, যেগুলো ঠিক নাও হতে পারে। সেটা নিয়েই নীচের অংকটা।

Exercise 32: System of linear equations নিয়ে চারটে বাক্য দিলাম। এরা ঠিক না ভুল বল।

1. যদি equation-এর সংখ্যা unknown-এর সংখ্যার সমান হয়, তবে সব সময়েই সমাধান করা সম্ভব।
2. যদি equation-এর সংখ্যা unknown-এর সংখ্যার চেয়ে কম হয়, তবে সব সময়েই অন্ততঃ একটা সমাধান থাকবে।
3. যদি equation-এর সংখ্যা unknown-এর সংখ্যার চেয়ে কম হয়, তবে কখনোই unique solution থাকা সম্ভব নয়।
4. যদি equation-এর সংখ্যা unknown-এর সংখ্যার চেয়ে বেশী হয়, তবে কোনো সমাধান পাওয়া যাবে না।

■

এবার একটা ঠকানো প্রশ্ন।

Exercise 33: এমন কোনো একখানা equation কি সম্ভব, যার মধ্যে তিনটে unknown, এবং equation-টার unique solution আছে? আমরা এখানে $\mathbb{R}$-এর মধ্যে কাজ করছি।
HINT: একটা কথা বলে না দিলে তুমি ভুল লাইনে চিন্তা করতে থাকবে। আমি কিন্তু equation-টাকে linear বলে দিই নি! ■

DAY 21
System of linear equations (part 2)

একটা system of linear equations সমাধান করতে পারা বিজ্ঞানের বহু শাখাতেই গুরুত্বপূর্ণ কাজ। এটা হল linear algebra-র সহজ প্রয়োগগুলোর অন্যতম। সহজ বললাম বটে, কিন্তু system-টা একটু বড় হলেই এরকম সমাধান হাতে করে ওঠা দুঃসাধ্য। তখন যে RREF-এর ডাক পড়ে, সেটা হয়তো আন্দাজ করে ফেলেছ। আজকে সেটাই আলোচনা করব।
আগেই বলেছি যে, যদি A একটা nonsingular matrix হয়, তবে $A\vec{x} = \vec{b}$-র unique solution হবে $A^{-1}\vec{b}$. সুতরাং এরকম system-এর সমাধান করার একটা কায়দা হল প্রথমে A^{-1} বার করে, তা দিয়ে $\vec{b}$-কে বাঁদিক থেকে গুণ করে দেওয়া। সেই নিয়েই নীচের অংকটা।

Example 21: Solve by matrix inversion method the following system of equations:

$$\begin{aligned} x + 2y + 3z &= 14 \\ 2x - y + 5z &= 15 \\ -3x + 2y + 4z &= 13 \end{aligned}$$

SOLUTION:

The system is $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 5 \\ -3 & 2 & 4 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 14 \\ 15 \\ 13 \end{bmatrix}.$$

এবার আমরা Gauss-Jordan elimination ব্যবহার করে A^{-1} বার করব। তার কায়দা তো আগেই শিখেছি, প্রথমে ডানদিকে একটা identity matrix লাগিয়ে নিয়ে RREF বার করা--

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & -1 & 5 & 0 & 1 & 0 \\ -3 & 2 & 4 & 0 & 0 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{14}{57} & \frac{2}{57} & -\frac{13}{57} \\ 0 & 1 & 0 & \frac{23}{57} & -\frac{13}{57} & -\frac{1}{57} \\ 0 & 0 & 1 & -\frac{1}{57} & \frac{8}{57} & \frac{5}{57} \end{array}\right]$$

এখানে আর মাঝের ধাপগুলো করে দেখাচ্ছি না।

Thus

$$A^{-1} = \frac{1}{57}\begin{bmatrix} 14 & 2 & -13 \\ 23 & -13 & -1 \\ -1 & 8 & 5 \end{bmatrix}.$$

So the required solution is

$$A^{-1}\vec{b} = \frac{1}{57}\begin{bmatrix} 14 & 2 & -13 \\ 23 & -13 & -1 \\ -1 & 8 & 5 \end{bmatrix}\begin{bmatrix} 14 \\ 15 \\ 13 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

এখানে আমরা A-র পাশে একটা identity matrix লাগিয়ে নিয়েছিলাম। ফলে প্রত্যেক ধাপে যখন row-গুলোর উপর basic operation করছিলাম, সেগুলোর প্রভাবে identity matrix-টা পরিবর্তিত হতে হতে A^{-1} হয়ে গেল। তার মানে পুরো কাজটার ফল হল $[A \quad I]$-কে বাঁদিক থেকে A^{-1} দিয়ে গুণ করা, যাতে শেষে আসে

$$A^{-1}[A \quad I] = [I \quad A^{-1}].$$

কিন্তু যদি আমরা $[A \quad I]$-র বদলে $[A \quad \vec{b}]$ নিয়ে একই কাজ করতাম, তবে পেতাম

$$A^{-1}[A \quad \vec{b}] = [I \quad A^{-1}\vec{b}].$$

এইভাবে এগোনোর দুটো সুবিধা--এক, $[A \quad I]$-এর চেয়ে $[A \quad \vec{b}]$ সাইজে ছোটো, সুতরাং প্রত্যেক ধাপে কম পরিশ্রম করতে হয়। আর দুই, এতে একবারেই $A^{-1}\vec{b}$ বেরিয়ে যায়, প্রথমে A^{-1} বার করে সেটা দিয়ে $\vec{b}$-কে গুণ করতে হয় না। এবার আমরা এই অংকটা সেভাবে করে দেখাচ্ছি--

বিকল্প পদ্ধতি

We start with the augmented matrix.

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 14 \\ 2 & -1 & 5 & 15 \\ -3 & 2 & 4 & 13 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

So the solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

■

Exercise 34: Solve $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -4 & 1 \\ 3 & 0 & -3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 5 \\ -1 \\ -3 \end{bmatrix}.$$

■

Augmented matrix-এর RREF বার করাটা একটা শক্তিশালী অস্ত্র। আমরা এ যাবৎ কেবল একটা ক্ষেত্রেই এর প্রয়োগ দেখলাম, যখন coefficient matrix-টা nonsingular. কিন্তু এটা দিয়ে যেকোনো system of linear equations-কেই কব্জা করা যায়। কায়দাটা অতি সহজ--পুরো RREF-টাকে তুমি ফের আবার আলাদা আলাদা equation হিসেবে ভেবে নাও। অমনি উত্তর পেয়ে যাবে। কয়েকটা উদাহরণ দেখলেই বুঝবে।

Example 22: যদি আমাদের system-টা হয় $x + y = 2$ এবং $2x + 2y = 5$, তবে RREF-টা হবে এইরকম--

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 5 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & \mathbf{1} \end{array}\right].$$

এ থেকে কী বলা যায়?

SOLUTION: এটাকে ফের equation হিসেবে লিখলে শেষ row-এর equation-টা হবে $0x + 0y = 1$. অমনি বুঝতে পারছ যে, inconsistent হবে।

In general, যদি খাড়া লাইনটার ডানদিকে কোনো leading 1 থাকে (মোটা করে দেখিয়েছি), তবে inconsistent হবে। সেটা না থাকলে, consistent হওয়া কেউ ঠেকাতে পারত না! ■

Example 23: যদি এইরকম হত, তবে?

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 3 & 4 & 7 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} \mathbf{1} & \mathbf{0} & 1 \\ \mathbf{0} & \mathbf{1} & 1 \\ 0 & 0 & 0 \end{array}\right].$$

SOLUTION: এখানে equation হিসেবে লিখে নিলে পাবে $x = 1$, $y = 1$ এবং আর $0x + 0y = 0$ (যেটা বলাই বাহুল্য)। ব্যস, unique solution পেয়েই গেলে!

In general, যদি লাইনের ডানদিকে কোনো leading 1 না থাকে, এবং RREF-টার nonzero row-গুলো খাড়া লাইনটার বাঁদিকে একটা identity matrix সৃষ্টি করে (মোটা করে দেখিয়েছি), তবে unique solution থাকবে, এবং সেটা পাওয়ার জন্য ফের nonzero row-গুলোর দিকে তাকাতে হবে, কিন্তু এবার লাইনটার ডানদিকে। ওখানে যা আছে, সেটাই solution হতে বাধ্য। ■

Example 24: যদি এরকম হয়, তবে কী বলবে?

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 4 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right].$$

SOLUTION: RREF-এর row-গুলোকে ফের equation আকারে লিখে নিলেই পাবে $x + y = 2$ আর $0x + 0y = 0$ (বলাই বাহুল্য)। তাই general solution হবে $\{(x, y) \ : \ x = 2 - y, \ y \in F\}$.

In general, যদি লাইনের ডানদিকে কোনো leading 1 না থাকে, এবং nonzero row-গুলো লাইনের বাঁদিকে identity matrix সৃষ্টি না করে, তবেই একাধিক solution থাকবে। ■

এই তিনটে অংক থেকে যা শিখলে, তার নির্যাসটা দিলাম Fig 4-এ।

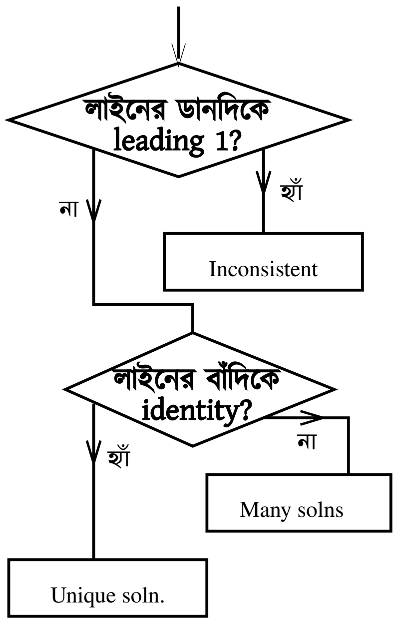


Fig 4

Exercise 35: নীচে প্রতি ক্ষেত্রে একটা করে augmented matrix-এর RREF দেওয়া আছে। তোমার কাজ হল general solution বার করা (বা inconsistent হলে, সেটা বলা)।

$$\textbf{(1)} \left[\begin{array}{ccc|c} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right] \quad \textbf{(2)} \left[\begin{array}{ccc|c} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \textbf{(3)} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

■

যদি একটা system of equations দিয়ে দেওয়া থাকে, $A\vec{x} = \vec{b}$, তবে তার solution আছে কিনা, বা থাকলে কটা আছে ইত্যাদি আমরা linear algebra না জেনেও ছোটোবেলায় শেখা কায়দা দিয়েই বার করে ফেলতে পারি। কিন্তু এই যে কায়দাগুলো শিখলাম, তা দিয়ে উল্টো দিকটাও করা যায়-- solution-টা কীরকম চাই, সেই অনুযায়ী A আর $\vec{b}$ সুবিধামত তৈরী করা যায়। প্রথমে একটা অংক দেখি যেখানে A পুরো দেওয়া আছে, এবং $\vec{b}$-এর কিছুটা দেওয়া আছে। আমাদের কাজ হবে $\vec{b}$-টাকে পুরো তৈরী করা।

Example 25: Find the values of a for which the system of equations given below is consistent. Also solve completely in each consistent case.

$$\begin{aligned} x+y+z &= 1, \\ 2x+3y-z &= a+1, \\ 2x+y+5z &= a^2+1. \end{aligned}$$

[5] **(2013.4a)**

SOLUTION: Augmented matrix-এর RREF বানানো শুরু করা যাক। মাঝের ধাপগুলো চেপে যাচ্ছি।

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 3 & -1 & a+1 \\ 2 & 1 & 5 & a^2+1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 4 & 2-a \\ 0 & 1 & -3 & a-1 \\ 0 & 0 & 0 & a^2+a-2 \end{array}\right]$$

এই অব্ধি এসে Gauss-Jordan elimination থমকে যাবে, কারণ ওই $a^2 + a - 2$-টা zero নাকি nonzero, সেটা না জানলে এগোনো যাচ্ছে না। যদি nonzero হয়, তবে ওটা একটা leading 1-এ পরিণত হবে। সর্বনাশ, খাড়া লাইনের ডানদিকে leading 1, তবে তো inconsistent হয়ে যাবে! অতএব $a^2 + a - 2 = 0$ হতেই হবে, মানে $a = -2$ বা $a = 1$ হতেই হবে। সেক্ষেত্রে শেষ row-টায় সব কিছুই শূন্য। তাই RREF-টা হবে--

$$\left[\begin{array}{ccc|c} 1 & 0 & 4 & 2-a \\ 0 & 1 & -3 & a-1 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

এবার এটাকে ফের আলাদা আলাদা equation হিসেবে লিখে ফেলা যাক--

$$\begin{aligned} x + 4z &= 2 - a \\ y - 3z &= a - 1. \end{aligned}$$

এখান থেকে general solution পাওয়া যাবে--

$$x = 2 - a - 4z, \quad y = a - 1 + 3z,$$

যেখানে z is any arbitrary scalar, এবং $a \in \{-2, 1\}$. ■

এখানে A-টা পুরো দেওয়া ছিল, $\vec{b}$-টার কিছু অংশ আমাদের বার করতে হচ্ছিল। এবার $\vec{b}$-টা পুরো দেওয়া থাকবে, আর আমাদের বার করতে হবে A-র কিছু অংশ।

Example 26: Determine the values of k so that the following system of equations has

1. no solutions,
2. more than only solution,
3. a unique solution:

$$\begin{aligned} x + y - z &= 1 \\ 2x + 3y + kz &= 3 \\ x + ky + 3z &= 2. \end{aligned}$$

SOLUTION: যথারীতি augmented matrix-টার RREF বার করার চেষ্টা করব। মাঝের ধাপগুলো আর দেখাচ্ছি না।

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 2 & 3 & k & 3 \\ 1 & k & 3 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -3-k & 0 \\ 0 & 1 & k+2 & 1 \\ 0 & 0 & (2-k)(3+k) & 2-k \end{array}\right]$$

এইবার আমাদের Gauss-Jordan elimination থমকাবে, কারণ ওই $(2-k)(3+k)$-টা শূন্য কিনা জানা নেই। তাই কেসে ভেঙে ভেঙে এগোতে হবে--

- যদি $k \notin \{-3, 2\}$ হয়, তবে ওইখানে একটা leading 1 এসে আমরা লাইনের বাঁদিকে identity পেয়ে যাব। মানে unique solution.
- যদি $k = 2$ হয়, তবে শেষের row-তে সব কিছুই শূন্য হয়ে যাবে, খাড়া লাইনের বাঁদিকে identity হবে না। এক্ষেত্রে একাধিক solution. যেহেতু $\mathbb{R}$-এর উপরে কাজ হচ্ছে, তাই infinitely many solutions.
- যদি $k = -3$ হয়, তবে তৃতীয় row-এর শেষ সংখ্যাটা nonzero হবে, তাই leading 1-এর জন্ম দেবে। আর লাইনের ডান দিকে leading 1 থাকলে কী হয়, সে তো জানোই--inconsistent!

■

A দেওয়া থাকলে $\vec{b}$ বার করা শিখেছি, $\vec{b}$ দেওয়া থাকলে A-ও বার করতে পারছি। কিন্তু যদি $A, \vec{b}$ কোনোটাই পুরো না দেওয়া থাকে? তাহলেও অনেক সময়ে $A, \vec{b}$ দুটোই বার করে দেওয়া যায়। এরকম একটা উদাহরণ দেখা যাক।

Example 27: Investigate for what values of λ, μ the following system of equations

$$\begin{aligned} x+y+z &= 6 \\ x+2y+3z &= 10 \\ x+2y+\lambda z &= \mu \end{aligned}$$

has

1. only one solution
2. no solution
3. infinitely many solutions.

SOLUTION: শুরু হবে augmented matrix-এর RREF বার করার চেষ্টা দিয়ে--

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 10 \\ 1 & 2 & \lambda & \mu \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & \lambda-3 & \mu-10 \end{array}\right]$$

এবার কেসে ভেঙে ভেঙে এগোতে হবে--

- যদি $\lambda \neq 3$ হয়, তবে ওই $\lambda - 3$-টার জায়গায় একটা leading 1 আসবে, ফলে লাইনের বাঁদিকে identity পেয়ে যাব। সেক্ষেত্রে unique solution.
- যদি $\lambda = 3$, কিন্তু $\mu \neq 10$, হয়, তবে leading 1-টা আসবে $\mu - 10$-এর জায়গায়, মানে লাইনের ডানদিকে। অমনি inconsistent!
- যদি $\lambda = 3$ এবং $\mu = 10$ হয়, তবে তৃতীয় row পুরোই শূন্য দিয়ে তৈরী। Nonzero row-গুলো দিয়ে লাইনের বাঁদিকে identity হচ্ছে না। তাই একাধিক solution.

■

এই অংকটা করতে গিয়ে হয়তো পরিশ্রম বাঁচানোর একটা পথ তোমার চোখে পড়েছে। আমরা প্রতিটা row-কে তার প্রথম nonzero entry দিয়ে ভাগ করে leading 1 বানাচ্ছি। সেই ভাগটা না করলেও চলত। Leading 1-এর উপরের entry-গুলোকে শূন্য করারও দরকার ছিল না। ওসব লাগে যখন আমরা system-টার solution বার করতে চাই। এখানে আমাদের উচ্চাশা কম, solution-এর সংখ্যা জানলেই আমরা খুশি।

একইরকম আরো কয়েকটা অংক দিলাম নীচে তোমার হাত পাকানোর জন্য।

Exercise 36: Investigate for what values of a, b the following system of equations

$$\begin{aligned} x+y+z &= 1 \\ x+2y-z &= b \\ 5x+7y+az &= b^2 \end{aligned}$$

has

1. only one solution
2. no solution
3. infinitely many solutions.

■

Exercise 37: For what values of a, b the following system of equations has

1. only one solution
2. no solution
3. infinitely many solutions

over the field of rational numbers? The equations are

$$\begin{aligned} x_1+4x_2+2x_3 &= 1 \\ 2x_1+7x_2+5x_3 &= 2b \\ 4x_1+ax_2+10x_3 &= 2b+1 \end{aligned}$$

■

Exercise 38: Consider the system of equations:

$$\begin{aligned} x+y+z &= b \\ 2x+y+3z &= b+1 \\ 5x+2y+az &= b^2. \end{aligned}$$

Determine the conditions for which the above system of equations has
(1) only one solution, **(2)** no solution, **(3)** many solutions.
■

Answers

1. $r(A_1)=1, r(A_2)=2$. **2.** $A=[\ \vec{e}_1 \quad \vec{e}_2 \quad \vec{e}_3 \quad \vec{0} \quad \vec{0}\ ]$ নিলেই হবে। **3.** A_k-র প্রথম k-খানা column নাও $\vec{e}_1, ..., \vec{e}_k$, বাকি column-গুলো $\vec{0}$ নাও। 10×4 চাইলে এই matrix-গুলোর transpose নিলেই চলবে। **4.** প্রথম k-খানা column নাও $\vec{e}_1, ..., \vec{e}_k \in F^m$, বাকি column-গুলো স্রেফ $\vec{0}$. **5.** **(1)** 1, **(2)** 3, **(3)** 2, **(4)** 2, মাঝের row-টা অন্য দুটো row-এর average (গড়)। **(5)** 0. **6.** হ্যাঁ। **7.**

> The given matrix is
> $$A=\begin{bmatrix}1&0&1\\0&1&0\\1&1&1\end{bmatrix}.$$
> The second row is not a multiple of the first. Also the third row is the sum of the first two.
>
> So $r(A)=2$.

8. এখানে RREF-টা হবে
$$\begin{bmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&0\\0&0&0&0\end{bmatrix}.$$
দেখতেই পাচ্ছ যে খালি তিনটে nonzero row পড়ে আছে, সুতরাং rank হবে 3. **9.** $\text{nullity}(A)=5-2=3$. **10.** (C). **11.** $\mathscr{C}(A)\subseteq F^m$, $\mathscr{R}(A),\mathscr{N}(A)\subseteq F^n$. **12.** যেহেতু 5-টা row-ই linearly independent, তাই $\dim(\mathscr{R}(A))=5$, মানে $r(A)=5$. সুতরাং $\text{nullity}(A)=7-r(A)=2$. **13.** $m>n$, কারণ এখানে $r(A)=n-\text{nullity}(A)=n$.এদিকে $r(A)\leq m$ হতে বাধ্য। **14.** **(1)** ঠিক, **(2)** ঠিক, **(3)** ভুল, $\mathscr{N}(A)=\{\vec{0}\}$, **(4)** ঠিক, **15.** $\text{nullity}(A)=20-r(A)=20-r(A')=20-(10-\text{nullity}(A'))=11$. **16.** A অবশ্যই square হবে। কারণ $A_{m\times n}$ হলে $r(A)+\text{nullity}(A)=n$ এবং $r(A')+\text{nullity}(A')=m$. যেহেতু $r(A)=r(A')$ আর বলেই দিয়েছে $\text{nullity}(A)=\text{nullity}(A')$, তাই $m=n$. এবার $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$ নিলেই বুঝবে যে $\text{nullity}(A)=0$ হওয়ার বা symmetric হওয়ার কোনো দরকার নেই। **17.** এখানে $\vec{v}=\begin{bmatrix}1\\i\end{bmatrix}\in\mathscr{R}(A)$. আবার $A\vec{v}=\vec{0}$-ও বটে, মানে $\vec{v}\in\mathscr{N}(A)$. তার মানে $\vec{v}\in\mathscr{R}(A)\cap\mathscr{N}(A)$. **18.** A-টা full column rank, কিন্তু full row rank নয়। B কিন্তু full row rank, যদিও full column rank নয়। C কোনোটাই নয়। D দুটোই। **19.** Full column rank হলে n, আর full row rank হলে m. **20.** অসম্ভব, দুদিকেই linearly independent set চাইলে matrix-টাকে square হতেই হবে! **21.** **(1)** ভুল, **(2)** ঠিক, **(3)** ঠিক, **(4)** ভুল, **(5)** ভুল, **(6)** ভুল। **22.** হ্যাঁ। **23.** 5, কারণ এখানে rank আর column-সংখ্যা সমান। **25.** আলবত্! **26.** ঠিক আগের theorem-টার প্রমাণের মতই এগোও। সেখানে পেয়েছিলে $\vec{y}'\vec{y}=0$. এখানে পাবে $\vec{y}^*\vec{y}=0$, মানে $|y_1|^2+\cdots+|y_m|^2=0$. সুতরাং আবার $\vec{y}=\vec{0}$ পেয়ে যাচ্ছ। বাকি অংশ আগের মতই। **27.** শুরুতে আছে $[A\quad I]=\left[\begin{array}{ccc|ccc}2&0&1&1&0&0\\3&3&0&0&1&0\\6&2&3&0&0&1\end{array}\right]$.
এর RREF হচ্ছে $\left[\begin{array}{ccc|ccc}1&0&0&\frac{3}{2}&\frac{1}{3}&\frac{-1}{2}\\0&1&0&\frac{-3}{2}&0&\frac{1}{2}\\0&0&1&-2&\frac{-2}{3}&1\end{array}\right]$. তাই $A^{-1}=\begin{bmatrix}\frac{3}{2}&\frac{1}{3}&-\frac{1}{2}\\-\frac{3}{2}&0&\frac{1}{2}\\-2&-\frac{2}{3}&1\end{bmatrix}$ **29.** i-th row-এর α গুণ j-th row-এর থেকে বিয়োগ করে দিলেই হবে।
মানে $(A(i,\alpha,j))^{-1}=A(i,-\alpha,j)$. **30.**

> We have $z=0$ from the first two equations. So we get $x-2y=0$, or $x=2y$. This also satisfies the last equation.
>
> Hence a vector is a solution of the given system iff it is of the form $(2y,y,0)=y(2,1,0)$.

> Thus, the solution space is $\text{span}\{(2,1,0)\}$.

তার মানে solution space-টা হল $(2,1,0)$ বরাবর লাইনটা।

> One possible basis is $\{(2,1,0)\}$. Dimension is 1.

31. **(1)** Unique. **(2)** একাধিক। **(3)** Inconsistent. **32.** তৃতীয় বাক্যটা শুধু সঠিক। **33.** হ্যাঁ, সম্ভব, যেমন $x^2+y^2+z^2=0$. এর unique solution হল $x=y=z=0$. **34.**

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 5 \\ 1 & -4 & 1 & -1 \\ 3 & 0 & -3 & -3 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

তাই উত্তর হল $x=1,\ y=1,\ z=2$. **35.** **(1)** নীচের row-টা থেকে বোঝা যাচ্ছে inconsistent. **(2)** $x\in\mathbb{R},\ y=1,\ z=2$. **(3)** $x=5,\ y=-1-2z,\ z\in\mathbb{R}$.
36. শুরু হবে RREF বার করার চেষ্টা করে--

> $$[A \quad \vec{b}] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & b \\ 5 & 7 & a & b^2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 3 & 2-b \\ 0 & 1 & -2 & b-1 \\ 0 & 0 & a-1 & b^2-2b-3 \end{array}\right]$$

এই পর্যন্ত এসে কেস ধরে ধরে এগোতে হবে--

> - If $a\neq 1$, then $r([A\ \vec{b}])=r(A)=3$. So the system has only one solution.
> - If $a=1$ but $b\notin\{-1,3\}$, then $r([A\ \vec{b}])=3>r(A)=2$. So the system has no solution.
> - If $a=1$ and $b\in\{-1,3\}$, then $r([A\ \vec{b}])=r(A)=2<3$. So the system has infinitely many solutions assuming that the field is infinite.

37. Row echelon form of augmented matrix: $\left[\begin{array}{ccc|c} 1 & 0 & 6 & 8b-7 \\ 0 & 1 & -1 & 2-2b \\ 0 & 0 & a-14 & 1-2b \end{array}\right]$

- only one solution if $a\neq 14$.
- no solution if $a=14$ but $b\neq\frac{1}{2}$.
- many solutions if $a=14$ and $b=\frac{1}{2}$.

38. Row echelon form of augmented matrix: $\left[\begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & b-1 \\ 0 & 0 & a-8 & b^2-2b-3 \end{array}\right]$

- only one solution if $a\neq 8$.
- no solution if $a=8$ but $b\notin\{-1,3\}$.
- many solutions if $a=8$ and $b\in\{-1,3\}$.

Chapter IV
Linear Transformation

DAY 22
গোড়ার কথা

আমরা এই অধ্যায়ে linear transformation নামে একটা জিনিস শিখব। এই ধারণাটা linear algebra-র মূল ধারণাদুটোর একটা। প্রথম মূল ধারণাটা ছিল vector space, আর তার পরেই স্থান হল এই linear transformation-এর।

22.1 জিনিসটা কী?

ধরো তোমাকে $f(x) = 2\sin x + 5\tan x$-কে differentiate করতে বললাম। তবে তুমি নিশ্চয়ই এভাবে এগোবে--

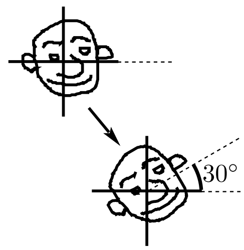
$$\begin{aligned}\frac{d}{dx}f(x) &= \frac{d}{dx}(2\sin x + 5\tan x)\\ &= 2\frac{d}{dx}\sin x + 5\frac{d}{dx}\tan x\\ &= 2\cos x + 5\sec^2 x.\end{aligned}$$

এখানে দ্বিতীয় ধাপটা লক্ষ কর যেখানে $\frac{d}{dx}$-টাকে ভেঙে দুভাগ করে দিলাম, এবং 2 আর 5 সংখ্যাদুটো $\frac{d}{dx}$-এর বাইরে বেরিয়ে এল। ঠিক একই ব্যাপার হয় integration-এর ক্ষেত্রেও, যেমন

$$\int af(x) + bg(x)dx = a\int f(x)dx + b\int g(x)dx.$$

বলাই বাহুল্য যে, এই ব্যাপারটা হয় বলে অংক কষা অনেক সহজ হয়। গোটা জিনিসটাকে নিয়ে একবারে কাজ না করে ভেঙে ভেঙে কাজ করা যায়। শুধু differentiation বা integration-ই নয়, এই ব্যাপারটা অন্য অনেক ক্ষেত্রেও দেখা যায়। যেমন ধরো Fig 1-এ একটা ছবিকে 30° ঘুরিয়ে দিয়েছি। এই কাজটাকে একটা function বলে ভাবা যায়--

Fig 1



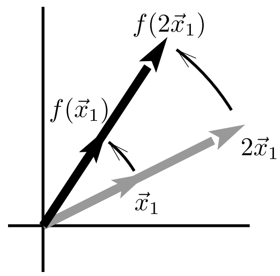


Fig 2

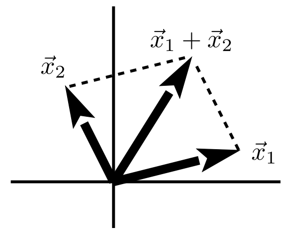


Fig 3

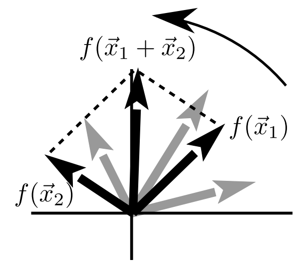


Fig 4

> কাগজটাকে মনে কর একটা গ্রাফ খাতা। যেই বিন্দুকে কেন্দ্র করে ঘোরাচ্ছ সেটাকে origin নাও। যেমন খুশী দুটো axis বসাও (Fig 1-এ একজোড়া axis দেখিয়েছি)। যদি কোনো বিন্দু $\vec{x}$ ঘোরানোর ফলে $\vec{y}$ বিন্দুতে উপস্থিত হয়, তবে বলব যে $f(\vec{x}) = \vec{y}$.

তাহলে লক্ষ কর যে $2\vec{x}$ অবশ্যই যাবে $2\vec{y}$-তে। Fig 2 দেখো। একইভাবে $f(3\vec{x}) = 3f(\vec{x})$, $f(-\vec{x}) = -f(\vec{x})$, ইত্যাদি হবে। মানে, কোনো সংখ্যা দিয়ে গুণ করলে সংখ্যাটা f-এর বাইরে বেরিয়ে আসছে। যদি দুটো বিন্দু নিই $\vec{x}_1$ আর $\vec{x}_2$ তবে $\vec{x}_1 + \vec{x}_2$-কে একটা parallelogram দিয়ে ভাবতে পারো Fig 3-এর মত। আশা করি সন্দেহ নেই যে 30° ঘোরালে সব কিছুই সমান তালে ঘুরবে। ফলে ঘোরানোর পরও parallelogram-টা parallelogram-ই থাকবে, অর্থাৎ $f(\vec{x}_1 + \vec{x}_2) = f(\vec{x}_1) + f(\vec{x}_2)$ হবে। Fig 4 দেখে নাও। সুতরাং সব মিলিয়ে দাঁড়ালো--যদি $\vec{x}_1, \vec{x}_2$ যেকোনো দুটো vector হয় এবং a, b যেকোনো দুটো scalar হয়, তবে

$$f(a\vec{x}_1 + b\vec{x}_2) = af(\vec{x}_1) + bf(\vec{x}_2)$$

হবে। এইধরণের function নিয়ে কাজ করা এত সুবিধা যে, লোকে ভালোবেসে এদের একটা নামই দিয়ে ফেলেছে--**linear transformation**. আমরা এইবার এদের কথাই শিখব। এতক্ষণ যা বললাম, সেটাকে আগে একটা সংজ্ঞার আকারে গুছিয়ে লিখে রাখি।

> DEFINITION: **Linear transformation**
>
> Let V, W be two vector spaces, both over the same field F. Then a function $f : V \to W$ is called a **linear transformation** if
>
> $$\forall \vec{u}, \vec{v} \in V \quad \forall a, b \in F \quad f(a\vec{u} + b\vec{v}) = af(\vec{u}) + bf(\vec{v}).$$

ছবি ঘোরানোর যে উদাহরণটা নিয়েছিলাম, সেখানে $V = W = \mathbb{R}^2$ ছিল, যাকে আমরা বোঝার সুবিধার জন্য গ্রাফ কাগজ বলে ভেবেছিলাম। কম্পিউটারের পর্দায় যখন কোনো ছবিকে ঘোরাতে হয়, তখন তা linear transformation-এর সাহায্যেই

Fig 5

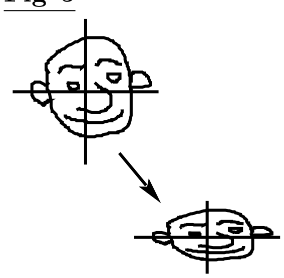
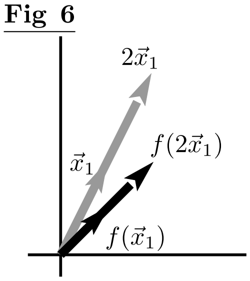


Fig 6



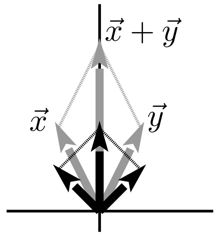


Fig 7

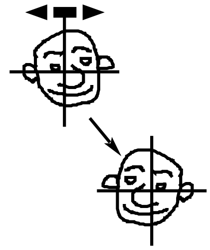


Fig 8

করা হয়। শুধু ঘোরানো কেন, আরও নানারকমভাবে ছবি নিয়ে নাড়াচাড়া করা যায় বিভিন্ন linear transformation দিয়ে। যেমন ধরো, যদি ছবিটাকে vertically চ্যাপ্টা করে অর্ধেক করে দিতে চাও Fig 5-এর মত, সেটাও হবে একটা linear transformation. কেন, সেটা ছবি দিয়ে চিন্তা করে দ্যাখো, একটা কোনো $\vec{x}$ নাও, সেটা যেন একটা তীর চিহ্ন, তাহলে $2\vec{x}$ হল সেই তীরটাকেই দ্বিগুণ করে বাড়িয়ে দেওয়া। যদি $\vec{x}$-কে তুমি vertically বেঁটে করে দাও, তবে $2\vec{x}$-ও একইভাবে বেঁটে হয়ে যেতে বাধ্য (Fig 6), মানে $f(2\vec{x}) = 2f(\vec{x})$. একইভাবে $\vec{x}+\vec{y}$-কে একটা parallelogram দিয়ে ভাবো (Fig 7)।

Exercise 1: ধরো আমি y-axis বরাবর একটা আয়না বসিয়ে ছবিটাকে সেই আয়নায় reflect করিয়ে দিলাম। Fig 8 দ্যাখো। সেটা কি একটা linear transformation হবে? যদি আয়নাটা বসাতাম x-axis বরাবর? বা $x = y$ লাইন বরাবর (Fig 9)? $x = 1$ লাইন বরাবর Fig 10? অংক কষে নয়, খালি ছবি দিয়ে ভেবে উত্তর দাও। ■

$\mathbb{R}^2$ থেকে $\mathbb{R}^2$-তে কোনো function দিলে এরকম ছবি দিয়ে ভাবার সুবিধা আছে। কিন্তু যদি $\mathbb{R}^3, \mathbb{R}^4$ ইত্যাদি থাকে, তবে সে সুবিধাটা নেই। তখন অংক কষে এগোনো ছাড়া পথ নেই। যেমন নীচের অংকটায়।

Example 1: Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$f(x_1, x_2, x_3) = (x_1 + x_2,\ \ x_2 + x_3,\ \ x_3 + x_1)$$

where $x_1, x_2, x_3 \in \mathbb{R}$. Show that f is a linear map.

SOLUTION: এখানে বলে রাখি যে linear map, linear mapping, linear function, linear transformation সবই একই জিনিসের বিভিন্ন নাম।

Shall show

$\forall \vec{x}, \vec{y} \in \mathbb{R}^3 \quad \forall a, b \in \mathbb{R} \quad f(a\vec{x} + b\vec{y}) = af(\vec{x}) + bf(\vec{y})$.

[$\forall \vec{x}, \vec{y}$] Take any $\vec{x} = (x_1, x_2, x_3), \vec{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$.

[$\forall a, b$] Take any $a, b \in \mathbb{R}$.

Fig 9

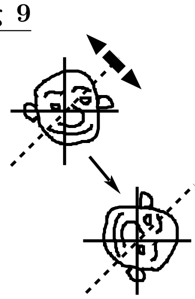
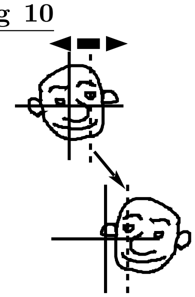


Fig 10



Then

$$\begin{aligned}
f(a\vec{x}+b\vec{y}) &= f(ax_1+by_1, ax_2+by_2, ax_3+by_3) \\
&\quad \text{এবার } f\text{-এর সংজ্ঞাটা লাগাব--} \\
&= (ax_1+by_1+ax_2+by_2, \\
&\quad ax_2+by_2+ax_3+by_3, \\
&\quad ax_3+by_3+ax_1+by_1) \\
&\quad x_i \text{ আর } y_i\text{-দের আলাদা করে লিখলে হয়--} \\
&= a(x_1+x_2,\ x_2+x_3,\ x_3+x_1)+ \\
&\quad b(y_1+y_2,\ y_2+y_3,\ y_3+y_1) \\
&= af(x_1,x_2,x_3)+bf(y_1,y_2,y_3),
\end{aligned}$$

as required.

■

একই কায়দায় নীচের অংকটা করো।

Exercise 2: Show that the following mapping $T:\mathbb{R}^3\to\mathbb{R}^3$ defined by

$$T(x_1,x_2,x_3) = (2x_1+x_2+3x_3,\ 3x_1-x_2+x_3,\ -4x_1+3x_2+x_3),$$

$(x_1,x_2,x_3)\in\mathbb{R}^3$ is linear. ■

22.2 চোখে দেখে চেনা

এবার একটা খুব সহজ উদাহরণ নিয়ে সেটাকে খুঁটিয়ে খুঁটিয়ে দেখি linear ব্যাপারটা কোথা থেকে আসছে।

Example 2: এখানে আমরা $f:\mathbb{R}\to\mathbb{R}$ নেব $f(x)=3x$. এটা কি একটা linear transformation?

SOLUTION: অবশ্যই, কারণ যদি $x_1,x_2\in\mathbb{R}$ নিই, তবে

$$f(x_1+x_2)=3(x_1+x_2)=3x_1+3x_2.$$

সুতরাং f-কে যে দুইভাগে ভেঙে ফেলা যাচ্ছে, সেটা আসলে যোগের উপরে গুণের distributivity ছাড়া আর কিছুই নয়। একইভাবে যদি $a\in\mathbb{R}$ আর $x\in\mathbb{R}$ নিই, তবে

$$f(ax)=3(ax)=(3a)x=(a3)x=a(3x)=af(x).$$

সুতরাং গুণ করা সংখ্যারা যে f-এর বাইরে বেরিয়ে আসে, সেটা হল গুণের associativity আর commutativity-র ফল। ■

এবার আগের অংকদুটোর দিকে ভালো করে তাকিয়ে দেখো। দেখবে সেখানেও একই ব্যাপার হচ্ছিল। এবার একটা function নেব, যেটা linear transformation নয়।

Example 3: $f:\mathbb{R}\to\mathbb{R}$ যদি $f(x)=x^2$ হয়, তবে সেটা কি একটা linear transformation হবে?

SOLUTION: উত্তর হল--না। যেমন $x_1 = 2, x_2 = 3$ নিলে $f(x_1 + x_2) \neq f(x_1) + f(x_2)$. ■

একই যুক্তিতে $f(x) = x^3$ বা $f(x) = \sin x$ ইত্যাদিরাও যে linear transformation হবে না বুঝতেই পারছ। সুতরাং linear transformation-কে linear বলার একটা কারণ আন্দাজ করতে পারছ। যদি $f : \mathbb{R} \to \mathbb{R}$ একটা linear transformation হয় তবে f-এর গ্রাফ হবে একটা সরলরেখা। কিন্তু এইখানে একটা সমস্যা আছে। গ্রাফটা সরলরেখা হলেই যে function-টা একটা linear transformation হবে এমনটা কিন্তু নয়! যেমন ধরো নীচের উদাহরণটা।

Example 4: $f : \mathbb{R} \to \mathbb{R}$ দিলাম $f(x) = x + 1$. এর গ্রাফ অবশ্যই একটা সরলরেখা। এটা কি একটা linear transformation?
SOLUTION: না, কারণ $f(2x) \neq 2f(x)$. ■

তার মানে $f(x) = mx + c$ হলেই function-টা linear transformation নাও হতে পারে। একটু চিন্তা করলেই বুঝবে যে, এখানে সমস্যাটা হল ওই c-টাকে নিয়ে। ওটা যদি না থাকত, তবে $f(x) = mx$ অবশ্যই একটা linear transformation হত।

Exercise 3: কোনো অংক না কষে খালি চোখে দেখেই ধাঁ করে বলতে হবে এদের মধ্যে কারা linear transformation আর কারা নয়।

1. $f : \mathbb{R}^2 \to \mathbb{R}^2$ যেখানে $f(x_1, x_2) = (x_1^2 - 3x_2,\ x_2)$,
2. $f : \mathbb{R}^3 \to \mathbb{R}$ যেখানে $f(x_1, x_2, x_3) = x_1 - x_3$,
3. $f : \mathbb{R}^2 \to \mathbb{R}^2$ যেখানে $f(x_1, x_2) = (x_1 - 3x_2,\ x_2 - 2)$,

■

Example 5: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). Let $S : \mathbb{R}^2 \to \mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^2$ be linear mappings defined by $S(x, y) = (x+y, y)$ and $T(x, y) = (x, x + y),\ (x, y) \in \mathbb{R}^2$. Then

(A) $TS - ST = I_{\mathbb{R}^2}$

(B) $(TS - ST)^2 = I_{\mathbb{R}^2}$

(C) $(TS - ST)^3 = I_{\mathbb{R}^2}$

(D) None of the above.

(where $I_{\mathbb{R}^2}$ denotes the identity mapping on $\mathbb{R}^2$). **(2019)**
SOLUTION:এই অংকে দুটো notation ব্যবহার করেছে, যেগুলো প্রথমে বোঝা দরকার। প্রথমতঃ, $I_{\mathbb{R}^2}$ মানে হল $\mathbb{R}^2$-এর উপরে identity linear transformation, অর্থাৎ $I_{\mathbb{R}^2}(x, y) = (x, y)$. এই function-এর পেটে তুমি যা-ই ঢোকাবে, সেটাই বেরিয়ে আসবে। দ্বিতীয়তঃ, TS মানে হল $T \circ S$, অর্থাৎ T এবং S-এর function composition, প্রথমে S লাগানো হয়েছে, তারপর T (ডানদিক থেকে বাঁদিকে)। যেমন $TS(x, y) = T(x + y, y) = (x + y, x + 2y)$. একইভাবে $ST(x, y) = S(x, x + y) = (2x + y, x + y)$. সুতরাং--

> $(TS - ST)(x, y) = TS(x, y) - ST(x, y) = (x+y, x+2y) - (2x+y, x+y) = (-x, -y)$.

এটা যে $I_{\mathbb{R}^2}$ নয়, সেটা তো বুঝতেই পারছ, ওই মাইনাসটা এসেই সব মাটি করে দিয়েছে! কিন্তু যদি square করি, তবে মাইনাসটা চলে যাবে। এই square করার জায়গাটায় একটু সাবধান। $(TS - ST)^2$ মানে হল পরপর দুবার $TS - ST$ লাগানো, অর্থাৎ--

> $(TS - ST)^2(x, y) = (TS - ST)(TS - ST)(x, y) = (TS - ST)(-x, -y) = (-(-x), -(-y)) = (x, y) = I_{\mathbb{R}^2}(x, y)$.

যদি $f(x)$ কোনো function হয়, তবে $f^2(x)$ বলতে আমরা সাধারণতঃ $f(f(x))$ বুঝিয়ে থাকি। তবে দুঃখের কথা, এর ব্যতিক্রম আছে, যখন আমরা $\sin^2(x)$ লিখে তখন মোটেই $\sin(\sin(x))$ বোঝাই না, তখন এর মানে হল $(\sin(x))^2$.
যাই হোক, এখানে $(TS - ST)^2$ মানে হল পরপর দুবার $TS - ST$ লাগানো। সুতরাং $(TS - ST)^3$ হবে পরপর তিনবার $TS - ST$ লাগানো, অর্থাৎ $(TS - ST)^2$-এর উপর আরো একবার $TS - ST$ লাগানো। তাই মাইনাসটা ফের ফিরে আসবে। সুতরাং উত্তর হবে (B). ■

Exercise 4: If $S(x, y) = (y, -x)$ and $T(x, y) = (-y, x)$, then which of the following is/are true?

(A) $S^2 = I_{\mathbb{R}^2}$ (B) $T^2 = I_{\mathbb{R}^2}$ (C) $TS = I_{\mathbb{R}^2}$ (D) $ST = I_{\mathbb{R}^2}$

■

Example 6: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that $T(1, 2) = (2, 3)$, $T(0, 1) = (1, 4)$. Then $T(5, 6)$ is

(A) $(6, -1)$ (B) $(-6, 1)$ (C) $(-1, 6)$ (D) $(1, -6)$

(2021.CC12.1h)

SOLUTION: এখানে T বলা আছে দুই জায়গায়, $(1, 2)$-এ আর $(0, 1)$-এ। আমাদের কাজ হল $(5, 6)$-এ T-এর value বার করা। এর জন্য আমাদের চেষ্টা থাকবে $(5, 6)$-কে ওই $(1, 2)$ আর $(0, 1)$-এর linear combination হিসেবে প্রকাশ করার। কারণ যদি একবার এমন $a, b \in \mathbb{R}$ পেয়ে যাই, যাতে $(5, 6) = a(1, 2) + b(0, 1)$ হয়, তবে T যেহেতু একটা linear transformation, আমরা বলতে পারব $T(5, 6) = aT(1, 2) + bT(0, 1)$.
প্রশ্ন হল এরকম linear combination হিসেবে লেখা যাবে কিনা।

> Let us see if we can find $a, b \in \mathbb{R}$ such that $(5, 6) = a(1, 2) + b(0, 1)$,
> ie, $a = 5$ and $2a + b = 6$, ie, $b = -4$.
> Since T is a linear transformation, $T(5, 6) = aT(1, 2) + bT(0, 1) = 5 \times (2, 3) - 4 \times (1, 4) = (6, -1)$.
> So answer is (A).

■

Exercise 5: Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation such that $T(1, 0, 0) = (2, 3)$, $T(1, 1, 0) = (3, 3)$ and $T(1, 1, 1) = (0, 0)$. Then $T(1, 2, 3) =$

(A) $(6,5)$ (B) $(5,6)$ (C) $(-5,-6)$ (D) $(-6,-5)$

■

এবার একটা সামান্য ঠকানো অংক।

Exercise 6: Let $T:\mathbb{R}^3\rightarrow\mathbb{R}^2$ be a linear transformation such that $T(1,1,1)=(1,1)$, $T(1,1,2)=(3,3)$ and $T(2,2,1)=(0,0)$. Then $T(1,2,3)=$

(A) $(6,5)$

(B) $(5,6)$

(C) $(-5,-6)$

(D) Cannot be found based on the given information.

■

এবার একটা বৈশিষ্ট্যের কথা বলি, যেটা যে কোনো linear transformation-এরই আছে। সেটা হল zero vector-এর উপর linear transformation লাগালে উত্তরটাও সর্বদা zero vector-ই হতে বাধ্য!

THEOREM

If $f:V\rightarrow W$ is a linear transformation, then $f(\vec{0}_V)=\vec{0}_W$.

এখানে দুইরকম $\vec{0}$ রয়েছে, লক্ষ করেছো তো? একটা হল V-এর zero vector, যাকে $\vec{0}_V$ লিখেছি, আর অন্যটা $\vec{0}_W$, সেটা W-এর zero vector.

Proof: প্রমাণটা খুবই সহজ।

$$f(\vec{0}_V)=f(\vec{0}_V+\vec{0}_V)=f(\vec{0}_V)+f(\vec{0}_V).$$

এখানে $f(\vec{0}_V)$ হল W-এর একটা vector. ধরো এর নাম দিলাম $\vec{w}$.

Let $\vec{w}=f(\vec{0}_V)$.
Then we have $\vec{w}=\vec{w}+\vec{w}$.

এবার দুদিক থেকেই $\vec{w}$ বিয়োগ করে দেব--

$\therefore\vec{w}=\vec{0}_W$, ie, $f(\vec{0}_V)=\vec{0}_W$, as required.

[Q.E.D]

Exercise 7: এমন একটা $f:\mathbb{R}^2\rightarrow\mathbb{R}$ দিতে পারো, যাতে $f(0,0)=0$ হয়, অথচ তাও f একটা linear transformation না হয়? ■

Exercise 8: নীচের দুটো বাক্যের মধ্যে একটাই খালি সঠিক। কোনটা?

1. $f(\vec{0}_V) = \vec{0}_W$ is a sufficient condition for f to be a linear transformation.
2. $f(\vec{0}_V) = \vec{0}_W$ is a necessary condition for f to be a linear transformation.

■

22.3 Infinite dimension-এর স্বাদ

আমরা জানি যে, vector space-রা দুইধরণের হয়--finite dimensional আর infinite dimensional. এই বইতে আমরা মূলতঃ finite dimensional-দের নিয়েই কাজ করছি। Linear transformation-দের বেলাতেও একই ব্যাপার। কিন্তু যখন বলছি যে $f : V \to W$ হল একটা linear transformation, তখন V এবং W দুজনেই বা অন্ততঃ একজন infinite dimensional হতেই পারে। এই অধ্যায়ে যা যা শিখব, তার অনেক কিছুই dimension-টা finite নাকি infinite, তার উপর নির্ভর করে না, দুইক্ষেত্রেই খাটে। কিন্তু যদি V, W দুজনেই finite dimensional হয়, তবে একটা বাড়তি সুবিধা হয়। সেক্ষেত্রে linear transformation-টাকে একটা matrix দিয়ে লেখা যায়। কাল আর পরশু আমরা এই ব্যাপারটা নিয়ে আলোচনা করব। আপাততঃ একটা ছোট্টো উদাহরণ দিই, যেখানে V হল infinite dimensional, কিন্তু W হল finite dimensional.

Example 7: আমরা V নেব $\mathbb{R}$ থেকে $\mathbb{R}$-এ যত রকম function হয়, তাদের set-টা। দুটো এরকম function-কে যোগ করলে আবার আরেকটা $\mathbb{R}$ থেকে $\mathbb{R}$-এ function পাওয়া যায়। আবার এরকম একটা function-কে কোনো real number দিয়ে গুণ করলেও একটা $\mathbb{R}$ থেকে $\mathbb{R}$-এ function পাওয়া যায়। সুতরাং V একটা vector space. আর W নেব $\mathbb{R}^2$. এবার একটা function নাও $T : V \to W$, এইরকম--

$$T(f) = (f(0), f(1)).$$

দেখাও যে T হল একটা linear transformation.

SOLUTION: প্রথমে বুঝে নিই T-টা ঠিক কী জিনিস। ধরো একটা function নিলাম $f : \mathbb{R} \to \mathbb{R}$ এইভাবে, $f(x) = e^x$. এইটা হল V-এর একটা vector. এর উপর আমরা T লাগাব, মানে $T(f)$ বার করব। বলা আছে $T(f) = (f(0), f(1))$. এখানে $f(0) = e^0 = 1$ আর $f(1) = e^1 = e$. সুতরাং $T(f) = (1, e)$.
যদি $2f$ নিয়ে কাজ করতাম (মানে যদি function-টা হত $2e^x$), তবে একইভাবে পেতাম

$$T(2f) = (2f(0), 2f(1)) = 2(f(0), f(1)).$$

দেখতেই পাচ্ছ যে, কোনো real number দিয়ে গুণ করলে সেটা T-এর বাইরে বেরিয়ে আসবে। একইভাবে যদি দুটো function নিতাম $f, g : \mathbb{R} \to \mathbb{R}$ তাদের যোগ করলে যে $h = f + g$ পাওয়া যায়, তার উপরে T লাগালে

$$\begin{aligned} T(h) &= (h(0), h(1)) = (f(0) + g(0), f(1) + g(1)) \\ &= (f(0), f(1)) + (g(0), g(1)) = T(f) + T(g). \end{aligned}$$

সুতরাং T একটা linear transformation. এখানে W-এর dimension হল 2. কিন্তু V-এর dimension হল infinite. সেটা বোঝা যায় এইভাবে। V-এর মধ্যে এই function-গুলো লক্ষ কর

$$f_n(x) = \begin{cases} 1 & \text{if } x = n \\ 0 & \text{otherwise} \end{cases}$$

যেখানে $n = 1, 2, 3, ...$ এইভাবে তুমি অসংখ্য function পেলে, $f_1, f_2, f_3, ...$ ইত্যাদি। এরা সবাই linearly independent. (কেন বল তো?) যেহেতু infinite সাইজের একটা linearly independent set পাওয়া যাচ্ছে, সুতরাং V আর finite dimensional হয় কী করে? ■

Example 8: এবার এমন একটা linear transformation দেখাই $T : V \to W$, যেখানে V আর W দুজনেই infinite dimensional. আমরা V নেব যাবতীয় $\mathbb{R}$ থেকে $\mathbb{R}$-এ differentiable function-এর vector space. আর W নেব $\mathbb{R}$ থেকে $\mathbb{R}$-এ যাবতীয় function-এর space. এখানে T নেব $differentiation$. যেমন $T(\sin(x)) = \cos x$. ■

DAY 23
লিখে প্রকাশ করা (theory)

অংকের জগতে আমরা নানান abstract ধারণা নিয়ে কাজ করি। এইরকম ধারণাকে মাথায় রাখা সহজ, খাতার পাতায় প্রকাশ করতে গেলে লেখার একটা কায়দা চাই। একই ধারণাকে অনেকভাবে লিখে বোঝানো যায়, তাদেরকে বলে একেকটা representation. যেমন যদি বলি একটা জিনিসকে সমান দুইভাগে ভাগ করলে প্রতি ভাগে কতটা করে পড়ে? উত্তর হল অর্ধেকটা করে। এই "অর্ধেক" ব্যাপারটা হল একটা ধারণা। এটাকে লিখে বোঝানো যায় নানাভাবে, যেমন fraction-এর (ভগ্নাংশের) আকারে, $\frac{1}{2}$ বা $\frac{2}{4}$ ইত্যাদি, কিংবা decimal (দশমিক) দিয়ে যেমন 0.5, কিংবা percent (শতকরা) আকারে, 50%. এরা সবাই দেখতে আলাদা হলেও একই ধারণার বিভিন্ন representation. একেকটা representation-এর একেকরকমের সুবিধা থাকে, তাই কখনো একভাবে লিখে নিলে সুবিধা হয়, কখনো বা অন্যভাবে। সেই কারণে একটা representation থেকে অন্য representation-এ যাওয়াটা শিখে রাখা ভালো। যেমন, ছোটোবেলায় শিখেছিলে fraction থেকে decimal, আর decimal থেকে fraction বার করতে।
এখানে যে ধারণাটা নিয়ে কাজ করছি সেটা হল linear transformation. এর নানা রকম representation এবার শিখব।

23.1 ভাষায় বর্ণনা করে

অনেক linear transformation আছে, যাদেরকে মুখের ভাষাতেই চমৎকার বর্ণনা করা চলে, যেমন Fig 11-এ যে linear transformation-টা দেখিয়েছি, সেটাকে বলা চলে "origin-কে কেন্দ্র করে ঘড়ির কাঁটার মত 30° ঘোরানো"। তেমনি Fig 12-এর বেলায় বলা যায়, "x-axis বরাবর চেপ্টে অর্ধেক করে দিয়ে y-axis বরাবর টেনে দ্বিগুণ করে দেওয়া"। তবে সমস্যা হল এইরকম জ্যামিতিক ভাষায় বর্ণনা খালি $\mathbb{R}, \mathbb{R}^2$ বা বড়জোর $\mathbb{R}^3$ পর্যন্তই চলে। $\mathbb{R}^3$-তেও মাঝে মাঝে সমস্যা হয়। যেমন ধরো Fig 13(a)-এর cube-টাকে ঘুরিয়ে Fig 13(b)-এর অবস্থানে আনা হয়েছে। চট্ করে এই ঘোরানোটাকে বর্ণনা করা শক্ত।

23.2 ফর্মুলা ব্যবহার করে

Linear transformation মাত্রেই একটা function, এবং অন্যান্য বহু function-এর মত এদেরকেও ফর্মুলা দিয়ে লেখা যায়। যেমন Fig 12-এর বেলায় $T : \mathbb{R}^2 \to \mathbb{R}^2$-কে আমরা ফর্মুলা দিয়ে লিখতে পারি এইভাবে--

$$T(x, y) = \left(\tfrac{1}{2}x, 2y\right).$$

কখনো কখনো এইভাবে ফর্মুলা দিয়ে লেখাকে বলে linear transformation-টার **explicit representation.**

Fig 11

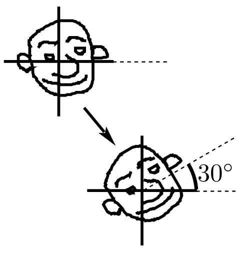


Fig 12

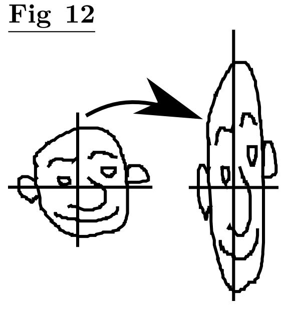
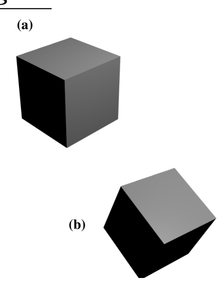


Fig 13



23.3 Basis-এর image দিয়ে

গতকাল আমরা একটা অংক করেছিলাম, যেখানে $T : \mathbb{R}^2 \to \mathbb{R}^2$ একটা linear transformation ছিল, এবং $T(1,2)$ আর $T(0,1)$ বলে দিয়ে $T(5,6)$ বার করতে দিয়েছিল। সমাধান করার জন্য আমরা $(5,6)$-কে $(1,2)$ আর $(0,1)$-এর linear combination হিসেবে লিখে নিয়েছিলাম--

$$(5,6) = 5 \times (1,2) - 4 \times (0,1),$$

এবং তা থেকে $T(5,6) = 5T(1,2) - 4T(0,1)$ লিখতে পেরেছিলাম। এখানে $(5,6)$-এর বদলে $\mathbb{R}^2$-র অন্য যেকোনো vector থাকলেও চলত, কারণ $\{(1,2),(0,1)\}$ হল $\mathbb{R}^2$-র একটা basis, তাই $\mathbb{R}^2$-র মধ্যে যেকোনো vector-কেই ওদের linear combination হিসেবে লেখা যাবে।

তার মানে যদি $T : V \to W$ একটা linear transformation হয়, আর $\{\vec{v}_1, ..., \vec{v}_n\}$ হয় V-এর যেকোনো একটা basis, তবে $T(\vec{v}_1), ..., T(\vec{v}_n)$ বলে দিলেই পুরো T বলে দেওয়া হল।

এই কথাটাকেই গুছিয়ে লিখলে এই theorem-টা হয়--

THEOREM

Let V and W be two vector spaces over a field F. Let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ be a basis of V. Let $\vec{w}_1, , \vec{w}_2, ..., \vec{w}_n$ be any vectors (not necessarily distinct) in W. Then there exists unique linear transformation $T : V \to W$ such that $T(\vec{v}_i) = \vec{w}_i$ for $i = 1, 2, ..., n$.

Proof:

<u>Existence:</u> Since $\{\vec{v}_1, ..., \vec{v}_n\}$ is a basis of V,

$$\therefore \forall \vec{v} \in V \quad \exists \text{ unique } a_1, ..., a_n \in F \quad \vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n.$$

We define a function $T : V \to W$ as

$$\forall \vec{v} \in V \quad T(\vec{v}) = a_1\vec{w}_1 + \cdots + a_n\vec{w}_n.$$

This is a well-defined function.

⟦ <u>Because:</u>

a_i's are uniquely determined by $\vec{v}$.

⟧

এবার দেখাব যে, এই T-টা একটা linear transformation. মানে

$$\forall \vec{x}, \vec{y} \in V \quad \forall a, b \in F \quad T(a\vec{x} + b\vec{y}) = aT(\vec{x}) + bT(\vec{y}).$$

Take any $\vec{x}, \vec{y} \in V$ and $a, b \in F$.

এবার $\vec{x}, \vec{y}$ দুজনকেই V-এর basis দিয়ে লিখব।

Let

$$\begin{aligned}\vec{x} &= p_1\vec{v}_1 + \cdots + p_n\vec{v}_n\\ \vec{y} &= q_1\vec{v}_1 + \cdots + q_n\vec{v}_n.\end{aligned}$$

So

$$\begin{aligned}T(\vec{x}) &= p_1T(\vec{v}_1) + \cdots + p_nT(\vec{v}_n) = p_1\vec{w}_1 + \cdots + p_n\vec{w}_n\\ T(\vec{y}) &= q_1T(\vec{v}_1) + \cdots + q_nT(\vec{v}_n) = q_1\vec{w}_1 + \cdots + q_n\vec{w}_n.\end{aligned}$$

Then

$$a\vec{x} + b\vec{y} = (ap_1 + bq_1)\vec{v}_1 + \cdots + (ap_n + bq_n)\vec{v}_n.$$

Hence

$$T(a\vec{x} + b\vec{y}) = (ap_1 + bq_1)\vec{w}_1 + \cdots + (ap_n + bq_n)\vec{w}_n = aT(\vec{x}) + bT(\vec{y}).$$

Thus T is a linear transformation.

আমরা একভাবে একটা linear transformation বানিয়েছি। এবার তার uniqueness দেখাব, মানে যেভাবেই linear transformation বানাও না কেন, theorem-টার শর্ত মানতে হলে তুমি আমাদের linear transformation-টাতে এসেই পড়বে। দেখানোর জন্য আমরা theorem-এর শর্ত মেনে যা খুশি একটা linear transformation নেব, ধরো S, এবং দেখাবো যে, $\forall \vec{v} \in V \quad S(\vec{v}) = T(\vec{v})$ হতে বাধ্য।

Uniqueness: Let $S : V \to W$ be a linear transformation such that

$$\forall i \quad S(\vec{v}_i) = \vec{w}_i.$$

Shall show that $\forall \vec{v} \in V \quad S(\vec{v}) = T(\vec{v})$.

Take any $\vec{v} \in V$.

Then we can write

$$\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n.$$

So

$$\begin{aligned}S(\vec{v}) &= S(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n)\\ &= a_1S(\vec{v}_1) + \cdots + a_nS(\vec{v}_n)\\ &= a_1\vec{w}_1 + \cdots + a_n\vec{w}_n\\ &= a_1T(\vec{v}_1) + \cdots + a_nT(\vec{v}_n)\\ &= T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n)\\ &= T(\vec{v}),\end{aligned}$$

as required.

[Q.E.D]

23.4 Matrix ব্যবহার করে

একটা linear transformation-কে লিখে বোঝানোর যতরকম কায়দা আছে, তার মধ্যে নিঃসন্দেহে সবচেয়ে বহুলব্যবহৃত হল matrix. বস্তুতঃ matrix-এর জনপ্রিয়তা এই কারণেই।
এক্ষুণি basis-এর image দিয়ে যে কায়দাটা শিখলাম, matrix-এর কায়দাটা আসলে সেটারই একটু গোছানো রূপ। ধাপে ধাপে উদাহরণ দিয়ে বোঝাই--

- ধরো $T : V \to W$ হল একটা linear transformation.

  উদাহরণস্বরূপ, আমরা V নেব যাবতীয় $a+bx+cx^2+dx^3$ চেহারার polynomial-এর vector space (যেখানে $a,b,c,d \in \mathbb{R}$), আর W নেব যাবতীয় $a+bx+cx^2$ জাতীয় polynomial-এর vector space (যেখানে $a,b,c \in \mathbb{R}$)। এদের মধ্যে linear transformation-টা নেব differentiation, যেমন $T(a+bx+cx^2+dx^3) = b+2cx+3dx^2$.

- আমাদের linear transformation-টাকে আমরা matrix দিয়ে প্রকাশ করতে চাই। এর জন্য একটাই প্রয়োজনীয় শর্ত, V আর W দুজনকেই finite dimensional হতে হবে।

  আমাদের উদাহরণে তা-ই আছে।

- প্রথমে V-এর একটা basis $\{\vec{v}_1, ..., \vec{v}_n\}$, আর W-এর একটা basis $\{\vec{w}_1, ..., \vec{w}_m\}$ নিতে হয়।

  আমাদের উদাহরণে ধরো V-এর জন্য আমরা basis নিলাম $\{1, x, x^2, x^3\}$ আর W-এর জন্য $\{1, x, x^2\}$.

- এর ফলে V আর W-এর সদস্যদের এককেটা tuple বলে ভাবা যায়।

  যেমন $a+bx+cx^2+dx^3 \in V$ হয়ে যাবে (a,b,c,d). আবার তার derivative-টা মানে $b+2cx+3dx^2 \in W$ হয়ে যাবে $(b, 2c, 3d)$.

- এবার আমরা এমন একটা matrix বানাব, A যাতে যদি যেকোনো $\vec{v} \in V$-এর tuple-টাকে A দিয়ে গুণ করে দিলে $T(\vec{v})$-এর tuple-টা পাওয়া যায়। অর্থাৎ linear transformation-টা পরিণত হয় একটা matrix দিয়ে গুণ করাতে।

  আমাদের উদাহরণে আমরা এমন A চাইব যাতে

  $$\begin{bmatrix} b \\ 2c \\ 3d \end{bmatrix} = A \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

  হয়। এখানে matrix-টা আন্দাজ করা কঠিন নয়--

  $$\begin{bmatrix} b \\ 2c \\ 3d \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}}_{A} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

- In general, এই matrix-টা বার করার কায়দা হল column ধরে ধরে এগোনো। ধরো, j-th column-টা বার করতে চাও। তার জন্য $\vec{v}_j$-এর উপর T লাগাও, এবং তারপর $T(\vec{v}_j)$-কে $\vec{w}_i$-গুলো দিয়ে লেখো--

  $$T(\vec{v}_j) = a_1\vec{w}_1 + \cdots + a_m\vec{w}_m.$$

  ব্যস্, এই $(a_1, ..., a_m)$-ই হবে matrix-টার j-th column.

উদাহরণস্বরূপ এই কায়দায় A-র তৃতীয় column-টা বার করে দেখাচ্ছি। V-এর basis-টা ছিল $\{1, x, x^2, x^3\}$, যার তৃতীয় সদস্য x^2-কে differentiate করে $2x$ পাব। এটাকে W-এর basis-এর সাপেক্ষে লিখলে হয় $(0, 2, 0)$. ঠিক এটাই ছিল matrix-টার তৃতীয় column.

Exercise 9: A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by

$$T(x_1, x_2, x_3) = (3x_1 - 2x_2 + x_3, x_1 - 3x_2 - 2x_3)$$

for all $(x_1, x_2, x_3) \in \mathbb{R}^3$. Find the matrix of T relative to the ordered bases $\big((0,1,0),(1,0,0),(0,0,1)\big)$ of $\mathbb{R}^3$ and $\big((0,1),(1,0)\big)$ of $\mathbb{R}^2$.[2] **(2020.CC6.3fii)**

HINT:
এখানে domain-এর dimension হল 3 আর codomain-এর 2. তাই matrix-টা হবে 2×3 সাইজের। এর প্রথম column-টা বার করে দেখাচ্ছি--

- $\mathbb{R}^3$-এর যে basis-টা দিয়েছে, তার প্রথম vector-টা নাও, $(0, 1, 0)$.
- এর উপর T লাগাও, $T(0, 1, 0) = (3 \times 0 - 2 \times 1 + 0, 0 - 3 \times 1 - 2 \times 0) = (-2, -3)$.
- এবার এটাকে $\mathbb{R}^2$-এর যে basis-টা দিয়েছে, তার সাপেক্ষে লেখো, $(-2, -3) = -3(0, 1) - 2(1, 0)$. এখানে coefficient-গুলো হল -3 আর -2. তাই প্রথম column-টা হবে $\begin{bmatrix} -3 \\ -2 \end{bmatrix}$.

একইভাবে বাকী column দুটো বার করো। ■

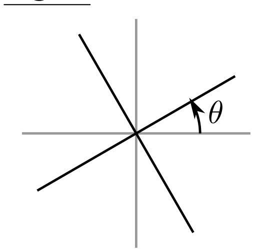
Exercise 10: A linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ is given by $T(x, y) = (x + y, y, y)$. Find the matrix representation of T with respect to the basis $B_1 = \{(1, 0), (0, 1)\}$ of $\mathbb{R}^2$ and $B_2 = \{(1, 1, 1), (0, 1, 0), (0, 0, 1)\}$ of $\mathbb{R}^3$.[5] **(2022.CC6.3d)**

HINT: আগের অংকের মত এখানে ordered basis বলে নি, কিন্তু সেটাই বুঝে নিতে হবে। ■

Exercise 11: এমন একটা 2×2 সাইজের matrix বানাও, যা দিয়ে গুণ করলে $\mathbb{R}^2$-টা origin-কে কেন্দ্র করে ঘড়ির কাঁটার উল্টোদিকে θ কোণে ঘুরে যায় (Fig 14)। এই matrix-টাকে বলে **Given's matrix**. ■

দুটো basis-এর সাপেক্ষে কোনো linear transformation-এর matrix বানানোর এই যে কায়দাটা শিখলাম, সেটা কেন কাজ করে, তা নিয়েই নীচের theorem-টা।

Fig 14



THEOREM

Let $T: V \to W$ be a linear transformation. Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be a basis of V, and let $\{\vec{w}_1, ..., \vec{w}_m\}$ be a basis of W. Then there exists a matrix $P_{m\times n}$ such that

$$T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) = b_1\vec{w}_1 + \cdots + b_m\vec{w}_m \tag{*}$$

if and only if

$$P\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}. \tag{**}$$

Proof: প্রথমে P-টা বানিয়ে নিই। এখানেই আগের অংকের কায়দাটা লাগবে--

> <u>Construction:</u> For each $j \in \{1, ..., n\}$ $T(\vec{v}_j) \in W$, and so
>
> $$T(\vec{v}_j) = p_{1j}\vec{w}_1 + \cdots + p_{mj}\vec{w}_m = \sum_i p_{ij}\vec{w}_i,$$
>
> for some p_{ij}'s.

এবার এই $p_{1j}, ..., p_{mj}$ দিয়ে P-এর j-th column-টা তৈরী করব, মানে--

> Let P be the matrix with (i, j)-th entry p_{ij}.

এইবার দেখাব যে, এই P দিয়ে সত্যিই আমাদের কাজ চলবে--

> <u>If part:</u>
>
> Let (**) hold, ie,
>
> $$b_i = \sum_j a_j p_{ij}.$$

Then

$$\begin{aligned} T(\vec{v}) &= T(a_1 v_1 + \cdots + a_m \vec{v}_m) \\ &= a_1 T(\vec{v}_1) + \cdots + a_m T(\vec{v}_m) \quad [\because T \text{ is linear}] \\ &= \sum_j a_j T(\vec{v}_j) \\ &= \sum_j \left[a_j \sum_i p_{ij} \vec{w}_i \right] \\ &= \sum_j \sum_i a_j p_{ij} \vec{w}_i \\ & \qquad \text{এবার sum দুটোর স্থানবিনিময় করব--} \\ &= \sum_i \sum_j a_j p_{ij} \vec{w}_i \quad \left[\text{interchanging sums} \right] \\ &= \sum_i \left[\vec{w}_i \underbrace{\sum_j a_j p_{ij}}_{b_i} \right] \\ &= \sum_i b_i \vec{w}_i, \end{aligned}$$

as required in (*).

এবার উল্টো দিকটা--

<u>Only if part:</u>

Let (*) hold.

Then $T(\vec{v}) = b_1 w_1 + \cdots + b_m \vec{w}_m$.

এবার $T(\vec{v})$-কে বড় করে লিখব, ঠিক এক্ষুণি যেমন করলাম--

> Now
>
> $$\begin{aligned} T(\vec{v}) &= T(a_1 v_1 + \cdots + a_m \vec{v}_m) \\ &= a_1 T(\vec{v}_1) + \cdots + a_m T(\vec{v}_m) \\ &= \sum_j a_j T(\vec{v}_j) \\ &= \sum_j \left[a_j \sum_i p_{ij} \vec{w}_i \right] \\ &= \sum_j \sum_i a_j p_{ij} \vec{w}_i \\ &\quad \text{এবার sum-দুটোর স্থানবিনিময় করব--} \\ &= \sum_i \sum_j a_j p_{ij} \vec{w}_i \quad \left[\text{interchanging sums} \right] \\ &= \sum_i \left[\vec{w}_i \sum_j a_j p_{ij} \right] \\ &= \sum_i \left[\sum_j a_j p_{ij} \right] \vec{w}_i. \end{aligned}$$

তার মানে $T(\vec{v})$-কে আমরা দুইভাবে লিখতে পারলাম--এক, $\sum_j b_j \vec{w}_j$ হিসেবে, আর দুই, $\sum_i \left[\sum_j a_j p_{ij} \right] \vec{w}_i$ হিসেবে। সুতরাং এই দুটো জিনিস নিশ্চয়ই সমান--

> Thus
>
> $$\begin{aligned} \sum_i \left[\sum_j a_j p_{ij} \right] \vec{w}_i &= \sum_i b_i \vec{w}_i \\ \text{or, } \sum_i \left[b_i - \sum_j a_j p_{ij} \right] \vec{w}_i &= \vec{0}. \end{aligned}$$

যেহেতু $\vec{w}_1, ..., \vec{w}_m$-গুলো linearly independent, তাই প্রত্যেকটা coefficient আলাদা করে 0 হবে, মানে--

> Since $\vec{w}_1, ..., \vec{w}_m$ are linearly independent, hence
>
> $$\forall i \in \{1, ..., m\} \quad b_i = \sum_j a_j p_{ij},$$
>
> as required in (**).

[Q.E.D]

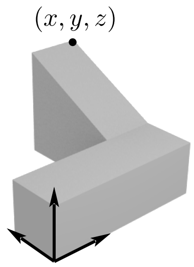


Fig 15

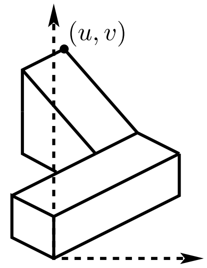


Fig 16

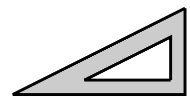


Fig 17

এখানে একটা বিষয় লক্ষণীয়-- যদি তোমাকে দুটো finite dimensional vector space দিই V আর W, এবং তাদের মধ্যে কোনো linear transformation $f : V \to W$ দিয়ে "f-এর matrix-টা" বার করতে বলি, তবে কিন্তু তার কোনো মানে হয় না। এখানে "f-এর matrix-টা" বলে কোনো কথা নেই, তুমি V এবং W-এর কী কী basis নেবে তার উপর নির্ভর করে একই f-এর বিভিন্ন matrix পাবে।

23.4.1 জ্যামিতি বাক্সে set-square কেন থাকে?

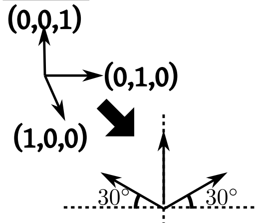
ছোটোবেলায় যখন প্রথম জ্যামিতি বাক্স হাতে পেয়েছিলে, তখন নিশ্চয়ই প্রশ্ন জেগেছিল--এই সব যন্ত্রপাতি দিয়ে কী হয়? স্কেল, কম্পাস ইত্যাদির ক্ষেত্রে এই প্রশ্নের সদুত্তর পেতে অসুবিধা হয় না। কিন্তু ওই ত্রিভুজাকৃতি set-square দুটো দিয়ে কী হয় সেই প্রশ্নের উত্তর আর পাওয়া হয় না। আসলে ওদের উদ্দেশ্য হল এক ধরণের linear transformation বার করা। খুলে বলি। ইঞ্জিনিয়াররা যখন কোনো নতুন জিনিস বানান তখন জিনিসটা কিরকম হতে যাচ্ছে, সেটার মাপজোখ ইত্যাদি প্রথমে মনে মনে ছকে নেন, কতকটা Fig 15-এর মত। যে মিস্ত্রী জিনিসটা বানাবেন তাঁর অবশ্য অত কল্পনাশক্তি নাও থাকতে পারে। তাই ইঞ্জিনিয়ার মনশ্চক্ষে যেটা কল্পনা করেছেন, সেটাকে Fig 16-এর মত দস্তুরমত কাগজে এঁকে দিতে না পারলে সেই কল্পনাকে বাস্তবায়িত করা যাবে না। Fig 15-এর জিনিসটা 3-dimensional, সুতরাং এর প্রতিটা বিন্দুকে (x, y, z) আকারে লেখা যায়, এবং ইঞ্জিনিয়ারের মনে প্রতিটা বিন্দুর (x, y, z) সম্বন্ধে পরিষ্কার ধারণা আছে। সেখান থেকে তাঁকে পৌঁছতে হবে Fig 16-এর মত খাতার পাতায় আঁকা একটা 2-dimensional ছবিতে। যদি খাতার পাতাটাকে গ্রাফ কাগজ ভেবে নাও, তবে তার দুটো axis-কে u আর v নাম দিলে ছবির প্রতিটা বিন্দু হবে (u, v). সুতরাং ইঞ্জিনিয়ার যখন তাঁর কল্পনাকে খাতার পাতায় ফুটিয়ে তুলবেন, তখন তাঁকে প্রতিটা (x, y, z)-এর জন্য (u, v) বার করতে হবে। বুঝতেই পারছ যে, এটা $\mathbb{R}^3$ থেকে $\mathbb{R}^2$-তে একটা function. এটা আসলে একটা linear transformation!

3-dimensional জিনিসটাকে কোন দিক থেকে দেখে আঁকছি, তার উপর নির্ভর করে linear transformation-টা নানারকম হতে পারে। জ্যামিতি বাক্সের দুটো set square এরকম দুটো linear transformation-এর জন্য। ওরা আসলে ছদ্মবেশে একেকটা matrix!

Fig 17-এ যে set square-টা দেখিয়েছি সেটার জন্য linear transformation-টা দেব basis-এর image ব্যবহার করে। Fig 18-এর দিকে চোখ রাখো। $(1, 0, 0)$ vector-টাকে আঁকা হবে horizontal-এর সঙ্গে 30° কোণ করে, মানে

$$f(1, 0, 0) = \left(\tfrac{\sqrt{3}}{2}, \tfrac{1}{2}\right).$$

Fig 18



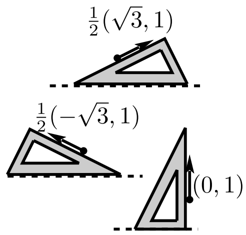


Fig 19

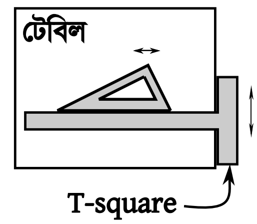


Fig 20

$(0,1,0)$-কে আঁকা হবে তার উল্টো দিকে $30°$ কোণে, মানে

$$f(0,1,0) = \left(\frac{-\sqrt{3}}{2}, \frac{1}{2}\right),$$

আর $(0,0,1)$ হবে খাড়া উপর দিকে, অর্থাৎ

$$f(0,0,1) = (0,1).$$

তার মানে যদি $\mathbb{R}^3$-এর basis নিই $\{(1,0,0),(0,1,0),(0,0,1)\}$ এবং $\mathbb{R}^2$-র basis নিই $\{(1,0),(0,1)\}$ তবে matrix-টা হল

$$P = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}.$$

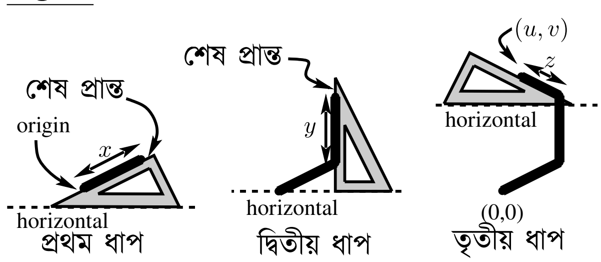
এই matrix-টাই যে ছদ্মবেশে Fig 17-এর set-square-টা, সেটা কি ঘুণাক্ষরেও সন্দেহ করেছিলে? সেটা বোঝার জন্য Fig 19-এর দিকে তাকাও। Matrix-টার তিনটে column পাওয়ার জন্য set-square-টাকে তিনভাবে ধরতে হয়, এবং ওর গা বরাবর যেমন দেখিয়েছি সেভাবে তীর চিহ্ন আঁকতে হবে। তীর চিহ্নগুলোর দৈর্ঘ্য সবক্ষেত্রেই 1, যেটা তুমি set-square-এর গায়ের স্কেল থেকেই মেপে নিতে পারবে। প্রতি ক্ষেত্রেই set-square-এর নীচের বাহুটাকে horizontal রাখতে হয়। খালি হাতে সেটা নিখুঁতভাবে করা মুস্কিল, তাই ইঞ্জিনিয়াররা তার জন্য T-square বলে আরেকটা জিনিস ব্যবহার করেন (Fig 20)। এর মাথাটাকে টেবিলের ধার বরাবর চেপে সরালে ডাণ্ডাটা সর্বদা horizontal থাকে, আর set-square-টাকে সেই ডাণ্ডা বরাবর সরানো হয়। এবার দ্যাখো (x,y,z) বিন্দুটার জন্য কী করে set-square ব্যবহার করে

$$\begin{bmatrix} u \\ v \end{bmatrix} = P \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

বার করা যাবে।
আমরা জানি

$$\begin{bmatrix} u \\ v \end{bmatrix} = P \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} + y \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} + z \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Fig 21



প্রথমে Fig 21-এর প্রথম ধাপের মত করে x মেপে নাও। তাহলে পেলে $x\begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$. এবার এর শেষ প্রান্ত থেকে দ্বিতীয় ধাপের মত করে y মেপে নাও। তাহলে পৌঁছনো গেল $x\begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} + y\begin{bmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$-তে। এবার তার শেষ প্রান্ত থেকে তৃতীয় ধাপে z পরিমাণ এগোও। ব্যস্ এভাবে যে বিন্দুতে পৌঁছলে সেটাই হবে (u, v).

DAY 24
লিখে প্রকাশ করা (practice)

একই linear transformation-কে যে কতরকমভাবে লেখা যায়, সে তো গতকাল শিখলাম। এই সব বিভিন্ন representation-এর মধ্যে একটা থেকে আরেকটায় অনুবাদ করতে পারলে অনেক সময়ে সুবিধা হয়। এবার সেই কায়দাকানুনগুলো দেখব। আজকের পড়াটা দেখতে ছোটো। কিন্তু এর মধ্যে অনেক অংক আছে তোমার নিজে নিজে করার জন্য।

একটা basis-এর image ব্যবহার করে কীভাবে একটা linear transformation-কে প্রকাশ করা যায় সেটা আমরা জানি। ব্যাপারটা ছিল অনেকটা "কান টানলে মাথা আসা"-র মত। Linear transformation-এর domain-টা যেন মাথাটা। Basis-টা যেন তার কান। Linear transformation-এর ফলে basis-টা কোথায় গিয়ে পড়ছে, সেটা জানলেই domain-এর প্রতিটা vector কে কোথায় যাচ্ছে, সব জানা হয়ে যায়।

Example 9: A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ transforms the vectors $(1, 0, 0), (1, 1, 0), (1, 1, 1)$ to the vectors $(1, 3, 2), (3, 4, 0)$ and $(2, 1, 3)$, respectively. Find T and the matrix representation of T relative to the standard ordered basis of $\mathbb{R}^3$.

SOLUTION: এখানে "Find T" মানে T-এর explicit representation চাইছে, অর্থাৎ ফর্মুলা দিয়ে প্রকাশ করতে বলছে। তারপর চেয়েছে standard basis-এর সাপেক্ষে matrix হিসেবে। যেহেতু এখানে domain আর codomain একই, তাই একটা basis-এই কাজ হয়ে যাচ্ছে। প্রথমে সেই basis-এর সাপেক্ষে T-কে প্রকাশ করি, তা থেকে ফর্মুলা আর matrix দুই-ই পেয়ে যাব।

আমাদের কাজ হবে T-এর প্রভাবে standard basis-টা কোথায় যায়, সেটা বার করা। যেহেতু এখানে T-কে একটা অন্য basis-এর সাপেক্ষে দিয়েছে, তাই প্রথমে standard basis-টাকে সেই basis দিয়ে লিখে নেব, standard basis-এর $(1, 0, 0)$ তো সেই basis-এও আছে। আর $(0, 1, 0) = (1, 1, 0) - (1, 0, 0)$ এবং $(0, 0, 1) = (1, 1, 1) - (1, 1, 0)$. সুতরাং standard basis-এর উপর T-এর প্রভাবটা বার করতে অসুবিধা নেই--

Here

$$\begin{aligned} T(1,0,0) &= (1,3,2), \\ T(0,1,0) &= T((1,1,0)-(1,0,0)) \\ &= T(1,1,0)-T(1,0,0) \quad [\because T \text{ is linear}] \\ &= (3,4,0)-(1,3,2) = (2,1,-2), \\ T(0,0,1) &= T((1,1,1)-(1,1,0)) \\ &= T(1,1,1)-T(1,1,0) \\ &= (2,1,3)-(3,4,0) = (-1,-3,3). \end{aligned}$$

So for any $(x, y, z) \in \mathbb{R}^3$,

$$\begin{aligned} T(x,y,z) &= T(x(1,0,0)+y(0,1,0)+z(0,0,1)) \\ &= xT(1,0,0)+yT(0,1,0)+zT(0,0,1) \\ &= x(1,3,2)+y(2,1,-2)+z(-1,-3,3) \\ &= (x+2y-z,\ \ 3x+y-3z,\ \ 2x-2y+3z), \end{aligned}$$

which is the required explicit representation of T.

এবার standard basis-টা T-এর প্রভাবে যে তিনটে vector-এ গিয়েছিল, তাদেরকে পাশাপাশি column-এর মত জুড়ে দিলেই প্রয়োজনীয় matrix-টা পেয়ে যাব--

So the matrix representation wrt the standard ordered basis is

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & -3 \\ 2 & -2 & 3 \end{bmatrix}.$$

■

Exercise 12: A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ transforms the vectors $(1,1,1), (0,1,1), (0,0,1)$ to the vectors $(3,0,2), (5,-1,3)$ and $(6,1,2)$, respectively. Find T and the matrix representation of T relative to the standard ordered basis of $\mathbb{R}^3$. ■

Example 10: Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$f(x_1,x_2,x_3) = (x_1+x_2,\ \ x_2+x_3,\ \ x_3+x_1)$$

where $x_1, x_2, x_3 \in \mathbb{R}$. Find the matrix associated with it with respect to the standard ordered basis of $\mathbb{R}^3$.

SOLUTION: যেহেতু এখানে standard basis নিয়ে কাজ করতে বলেছে, তাই এই অংকটা চোখে দেখেই করে দেওয়া যায়--

The standard ordered basis of $\mathbb{R}^3$ is $\{(1,0,0),\ (0,1,0),\ (0,0,1)\}$. We want a matrix A such that

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1+x_2 \\ x_2+x_3 \\ x_3+x_1 \end{bmatrix}.$$

By inspection,

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

is the required matrix.

■

Exercise 13: Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$f(x_1, x_2, x_3) = (3x_1 - x_2 + x_3,\ \ x_1 + x_2 - x_3,\ \ x_2 + x_3 - x_1)$$

where $x_1, x_2, x_3 \in \mathbb{R}$. Find the matrix associated with it with respect to the standard ordered basis of $\mathbb{R}^3$. ■

এখানে standard basis নিয়ে কাজ করতে দেওয়ায় জীবন খুব সহজ হয়ে গেছিল। যদি অন্য কোনো basis দিত, তবে linear transformation-এর ফলে basis-টা কোথায় যায়, এসব নিয়ে মাথা ঘামাতে হত। নীচের অংকটা সেরকম।

Exercise 14: Let T be a linear transformation which is defined as follows:

$$T(x, y, z) = (x - y,\ \ x + 2y,\ \ y + 3z).$$

Find the matrix representation of T with respect to the ordered bases

$$\{(1, 0, 0),\ \ (0, 1, 0),\ \ (0, 0, 1)\} \text{ and } \{(1, 1, 0),\ \ (1, 0, 1),\ \ (0, 1, 1)\}.$$

HINT: খালি প্রথম column-টা বার করে দেখাচ্ছি, বাকি দুটো তুমি নিজে করবে। যদিও এখানে domain আর codomain একই, তাও দুটোর জন্য দুটো আলাদা basis দিয়েছে। Domain-এর basis-এর প্রথম vector-টা $(1, 0, 0)$. এর উপর T লাগালে $T(1, 0, 0) = \cdots = (1, 1, 0)$. একে codomain-এর basis-এর সাপেক্ষে লিখলে হয়

$$(1, 1, 0) = 1 \times (1, 1, 0) + 0 \times (1, 0, 1) + 0 \times (0, 1, 1).$$

তার মানে coefficient-গুলো দিয়ে তৈরী tuple-টা হল $(1, 0, 0)$. সেটাই হবে matrix-টার প্রথম column. ■

Exercise 15: A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by

$$T(x, y, z) = (2x + 3y + 3z,\ \ 2x + y + 3z,\ \ 2x + 2y).$$

Determine the matrix T relative to the ordered basis $(2, 1, 1), (1, 2, 1), (1, 1, 2)$ of $\mathbb{R}^3$. ■

Exercise 16: Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$T(x_1, x_2, x_3) = (2x_1 + x_2 - x_3,\ \ x_2 + 4x_3).$$

Find the matrix of T relative to the ordered bases

$$\big\{(0, 1, 1),\ \ (1, 0, 1),\ \ (1, 1, 0)\big\} \text{ and } \big\{(0, 1),\ \ (1, 1)\big\}$$

of $\mathbb{R}^3$ and $\mathbb{R}^2$. ■

এতক্ষণ linear transformation-টা দেওয়া থাকছিল ফর্মুলা দিয়ে। নীচের অংকে সেটা দিয়েছে matrix দিয়ে। অবশ্য matrix থেকে ফর্মুলায় যেতে আর কতক্ষণ?

Exercise 17: The matrix representation of a linear mapping $T:\mathbb{R}^3\to\mathbb{R}^3$ is defined by

$$\begin{bmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

relative to the standard basis of $\mathbb{R}^3$. Find the explicit representation of T and matrix representation of T relative to the basis

$$\{(1,1,1),\ (0,1,1),\ (0,0,1)\}.$$

■

Exercise 18: The matrix representation of a linear transformation $T:\mathbb{R}^3\to\mathbb{R}^3$ is

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & -2 & 2 \\ 1 & 0 & -4 \end{bmatrix}$$

relative to the standard ordered basis of $\mathbb{R}^3$. Find the explicit representation of T relative to the standard basis, and matrix representation of T relative to the ordered basis $\{(1,0,1),\ (0,1,1),\ (1,1,0)\}$. ■

Exercise 19: The matrix representation of a linear transformation $T:\mathbb{R}^3\to\mathbb{R}^3$ is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -3 \\ 0 & 2 & -1 \end{bmatrix}$$

relative to the standard ordered basis of $\mathbb{R}^3$. Find the explicit representation of T and matrix representation of T relative to the ordered basis $\{(0,1,2),\ (-1,0,1),\ (2,1,1)\}$. ■

ফের একটা একইরকম অংক, খালি এবার matrix-টা অন্য একটা basis-এর সাপেক্ষে দিয়েছে।

Exercise 20: The matrix of a linear mapping $T:\mathbb{R}^3\to\mathbb{R}^3$ with respect to the ordered basis $(0,1,1),(1,0,1),(1,1,0)$ of $\mathbb{R}^3$ is given by

$$\begin{bmatrix} 0 & 3 & 0 \\ 2 & 3 & -2 \\ 2 & -1 & 2 \end{bmatrix}$$

Find T. Find also the matrix of T relative to the ordered basis $(2,1,1),(1,2,1),(1,1,2)$.

HINT:

প্রথমে T-কে ফর্মুলা দিয়ে লিখতে বলেছে, তাই standard basis দিয়ে প্রথমে লিখে নেব।

Let the given (ordered) basis be $B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

অর্থাৎ কিনা $\vec{v}_1 = (0,1,1)$, ইত্যাদি। এবার লক্ষ্য করো, $\frac{1}{2}(\vec{v}_1+\vec{v}_2+\vec{v}_3) = (1,1,1)$. এটা ব্যবহার করে standard basis-টাকে এটার সাপেক্ষে লিখে ফেলা সহজ--

We observe that

$$\begin{aligned}
\vec{e}_1 &= (1,0,0) = \tfrac{1}{2}(\vec{v}_1+\vec{v}_2+\vec{v}_3) - \vec{v}_1 = \tfrac{1}{2}(-\vec{v}_1+\vec{v}_2+\vec{v}_3)\\
\vec{e}_2 &= (0,1,0) = \tfrac{1}{2}(\vec{v}_1+\vec{v}_2+\vec{v}_3) - \vec{v}_2 = \tfrac{1}{2}(\vec{v}_1-v_2+\vec{v}_3)\\
\vec{e}_3 &= (0,0,1) = \tfrac{1}{2}(\vec{v}_1+\vec{v}_2+\vec{v}_3) - \vec{v}_3 = \tfrac{1}{2}(\vec{v}_1+\vec{v}_2-\vec{v}_3).
\end{aligned}$$

Hence

$$T(\vec{e}_1) = \frac{1}{2}\begin{bmatrix} 0 & 3 & 0\\ 2 & 3 & -2\\ 2 & -1 & 2\end{bmatrix}\begin{bmatrix} -1\\ 1\\ 1\end{bmatrix} = \frac{1}{2}\begin{bmatrix} 3\\ -1\\ -1\end{bmatrix} \text{ w.r.t. } B.$$

So wrt the standard basis,

$$T(\vec{e}_1) = \tfrac{1}{2}(3\vec{v}_1 - \vec{v}_2 - \vec{v}_3) = (-1,1,1).$$

Similarly, $T(\vec{e}_2) = (1,1,-3)$ and $T(\vec{e}_3) = (3,1,5)$.

Hence

$$\begin{aligned}
T(x,y,z) &= x(-1,1,1) + y(1,1,-3) + z(3,1,5)\\
&= (-x+y+3z, x+y+z, x-3y+5z).
\end{aligned}$$

ব্যস্, ফর্মুলা পেয়ে গেলে। এবার বাকি অংশ আগের অংকগুলোর মতই। ■

Exercise 21: The matrix of a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^2$ relative to the ordered bases $\big((0,1,1),(1,0,1),(1,1,0)\big)$ of $\mathbb{R}^3$ and $\big((1,0),(1,1)\big)$ of $\mathbb{R}^2$ is $\begin{bmatrix} 1 & 2 & 4\\ 2 & 1 & 0\end{bmatrix}$. Find the matrix of T relative to the ordered bases $\big((1,1,0),(1,0,1),(0,1,1)\big)$ of $\mathbb{R}^3$ and $\big((1,1),(0,1)\big)$ of $\mathbb{R}^2$.[5] **(2019.CC6.3g)** ■

Exercise 22: The matrix of a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^2$ relative to the ordered basis $(0,1,1)$, $(1,0,1)$, $(1,1,0)$ of $\mathbb{R}^3$ and $(1,0)$, $(1,1)$ of $\mathbb{R}^2$ is

$$\begin{bmatrix} 2 & 1 & 4\\ 1 & 2 & 0\end{bmatrix}.$$

Find the explicit representation of T and the matrix of T relative to the ordered basis $(1,0,1)$ $(1,1,0)$, $(0,1,1)$ of $\mathbb{R}^3$ and $(1,2)$, $(1,1)$ of $\mathbb{R}^2$.
■

DAY 25
Basis-এর পরিবর্তন

একটা finite dimensional vector space-এর মধ্যে একটা basis হচ্ছে একটা coordinate system-এর মত। এর সাহায্যে যেকোনো vector-কে একটা tuple আকারে লেখা যায়। যদি দুটো finite dimensional vector space-এর মধ্যে কোনো linear transformation থাকে, তবে domain আর codomain-এর basis-এর সাহায্যে সেটাকে একটা matrix দিয়ে প্রকাশ করা যায়। অনেক সময়ে কাজের সুবিধার জন্য basis-টাকে বদলাতে হয়, ফলে এইসব tuple এবং matrix-রাও বদলে যায়। Basis বদলালে এরা তার সঙ্গে তাল রেখে এরা কীভাবে বদলায়, তার হিসেব জানাটা খুব দরকারী। সেটাই আলোচনা করব আজকে।

25.1 Change of basis matrix

ধরো তোমাকে একটা vector space-এর দুটো basis দিলাম, $\mathcal{B}_1$ আর $\mathcal{B}_2$. এদের ব্যবহার করে আমরা একটা matrix বানানো শিখব, যাকে বলব $\mathcal{B}_1$ থেকে $\mathcal{B}_2$-তে যাওয়ার **change of basis matrix.** প্রথমে বলি জিনিসটা কী, এবং কী করে বার করতে হয়। তারপর বলব এর গুরুত্বটা কী।

Example 11: ধরো আমাদের vector space-টা হল সব $a+bx+cx^2$ জাতীয় polynomial-এর, যেখানে $a, b, c \in \mathbb{R}$.

এখানে field-টা $\mathbb{R}$. এর দুটো basis হল $\mathcal{B}_1 = \{1+x, 2, x+x^2\}$ এবং $\mathcal{B}_2 = \{x^2-1, x+3, 4\}$. প্রথমটা থেকে দ্বিতীয়টায় যাওয়ার change of basis matrix বার করো।

SOLUTION: কায়দাটা হল $\mathcal{B}_1$-এর vector-গুলোকে $\mathcal{B}_2$-র সাপেক্ষে tuple হিসেবে লেখা, এবং tuple-গুলোকে column-এর মত পাশাপাশি সাজিয়ে একটা matrix বানানো। ব্যস্! এখানে $\mathcal{B}_1$-এর প্রথম vector হল $1+x$. একে $\mathcal{B}_2$-র সাপেক্ষে লিখলে হয়--

$$1+x = 0 \times (x^2-1) + 1 \times (x+3) + \left(-\tfrac{1}{2}\right) \times 4.$$

Coefficient-গুলোকে tuple-এর মত সাজালে পাচ্ছি $\left(0, 1, -\frac{1}{2}\right)$. এটাই হবে change of basis matrix-টার প্রথম column. একইভাবে অন্য দুটো vector থেকে আসছে--

$$\begin{aligned} 2 &= 0 \times (x^2-1) + 0 \times (x+3) + \tfrac{1}{2} \times 4, \\ x+x^2 &= 1 \times (x^2-1) + 1 \times (x+3) + \left(-\tfrac{1}{2}\right) \times 4. \end{aligned}$$

সব মিলিয়ে change of basis matrix-টা দাঁড়াচ্ছে--

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

■

Exercise 23: আবার আগের অংকটাই করো। কিন্তু এবার change of basis matrix বার করো $\mathcal{B}_2$ থেকে $\mathcal{B}_1$-এ যাবার জন্য। আগের অংকের matrix-টার সঙ্গে এই অংকের matrix-টার কী সম্পর্ক হবে, আন্দাজ করতে পারছ? ■

Exercise 24: দুটো basis দিলাম $\mathbb{R}^3$-এর--

$$\mathcal{B}_1 = \{(1,2,3), (4,5,6), (7,8,10)\} \text{ এবং } \mathcal{B}_2 = \{(0,0,1), (0,1,1), (1,1,1)\}.$$

$\mathcal{B}_1$ থেকে $\mathcal{B}_2$-তে যাওয়ার change of basis matrix বার করো। ■

এবার দেখি change of basis matrix বার করে লাভটা কী হয়।

25.2 Tuple-এর উপর প্রভাব

Example 12: আবার আমরা সেই real vector space-টা নিয়েই কাজ করব, যার সদস্যরা হল সব $a + bx + cx^2$ জাতীয় polynomial, যেখানে $a, b, c \in \mathbb{R}$. সেই আগের দুটো basis-ই নেব, $\mathcal{B}_1 = \{1 + x, 2, x + x^2\}$ এবং $\mathcal{B}_2 = \{x^2 - 1, x + 3, 4\}$. যে vector-টা প্রথম basis-টার সাপেক্ষে $(3, 2, 1)$, সেটা দ্বিতীয় basis-টার সাপেক্ষে কত? এবার ধরো বলে দিলাম যে, প্রথমটা থেকে দ্বিতীয়টায় যাওয়ার জন্য change of basis matrix হল

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

এতে কি কিছু সুবিধা হবে?

SOLUTION: সহজবুদ্ধির কায়দাটা হল প্রথমে vector-টাকে বার করে নেওয়া, এবং তারপর সেটাকে দ্বিতীয় basis দিয়ে প্রকাশ করা। আমাদের vector-টা $\mathcal{B}_1$-এর সাপেক্ষে $(3, 2, 1)$, তার মানে vector-টা হল

$$3 \times (1 + x) + 2 \times 2 + 1 \times (x + x^2) = 7 + 4x + x^2.$$

এবার এটাকে $\mathcal{B}_2$ দিয়ে লিখব, মানে এমন $a, b, c \in \mathbb{R}$ খুঁজব, যাতে

$$7 + 4x + x^2 = a(x^2 - 1) + b(x + 3) + 4c$$

হয়। এখানে কাজটা চোখে দেখেই করা যাচ্ছে, $a = 1$, $b = 4$ এবং $c = -1$. সুতরাং $\mathcal{B}_2$-এর সাপেক্ষে আমাদের vector-টা হবে $(1, 4, -1)$.

Change of basis matrix ব্যবহার করে একই অংক অনেক সহজে হবে। এবার সেটা দেখাই।

<u>বিকল্প পদ্ধতি</u>

প্রথম basis-এর সাপেক্ষে যে tuple-টা আছে, সেটাকে change of basis matrix-টা দিয়ে গুণ করে দিলেই দ্বিতীয় basis-এর সাপেক্ষে tuple-টা পেয়ে যাবে--

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}.$$

■

দুটো কায়দার পার্থক্যটা একটা উপমা দিয়ে বোঝা যাক। মনে করো একটা দড়ি আছে, সেটা ইঞ্চিতে মেপে দেখলাম, 200 ইঞ্চি হল। বলো তো সেন্টিমিটারে মাপলে কী হত। এটা করার দুটো কায়দা--

- এক, ফের দড়িটা নাও, এবং নতুন করে সেন্টিমিটারে মেপে দেখো।
- দুই, ইঞ্চিতে মেপে যা পেয়েছিলে, সেটাকে 2.54 দিয়ে গুণ করে দাও।

আশা করি সন্দেহ নেই যে, দ্বিতীয় পথটা সহজতর। Change of basis matrix-টা এই 2.54-এর ভূমিকাটা পালন করছে। অবশ্য তুমি বলতেই পারো যে, change of basis matrix-টা বার করতেই তো অনেক খাটতে হয়েছিল। হ্যাঁ, কিন্তু সে তো ওই দুটো basis-এর জন্য খালি একবার, প্রতিটা tuple-এর পরিবর্তনের সময়ে সেই পরিশ্রমটা নতুন করে তো আর করতে হচ্ছে না! তবে সে পরিশ্রমটাও অনেকটাই কমানো যায়। সেটা একটু সংক্ষেপে বলে নিই।

25.2.1 সহজে বার করা

যেহেতু এখনকার সিলেবাসে change of basis matrix বার করাটা খুব গুরুত্বপূর্ণ স্থান অধিকার করে নেই, তাই পরীক্ষানম্বরপিপাসুরা এই অংশটুকু বাদ দিয়ে যেতে পারো।
সহজে change of basis matrix বার করার কায়দাটার পিছনে তিনটে ধারণা কাজ করছে। প্রথম ধারণাটা হল খালি F^n জাতীয় vector space-এর জন্য। দ্বিতীয় ধারণাটা হল RREF লাগানো, আর তৃতীয় ধারণাটা হল যে কোনো finite dimensional vector space-কেই F^n জাতীয় রূপে এনে ফেলা।
নীচের অংকটা প্রথম ধারণাটা নিয়ে।

Example 13: আমরা কাজ করব $\mathbb{R}^3$-র দুটো basis নিয়ে--

$$\mathcal{B}_1 = \{(1,2,3),(4,5,6),(7,8,10)\} \text{ এবং } \mathcal{B}_2 = \{(1,0,0),(1,1,0),(1,1,1)\}.$$

প্রথমটা থেকে দ্বিতীয়টাতে যাবার change of basis matrix বার করো।
SOLUTION: অংকটা গায়ের জোরে করার কায়দাটা আমরা আগেই শিখেছিলাম, যেটা যেকোনো vector space-এর জন্যই কাজ করে। কিন্তু এখানে যেহেতু vector space-টা F^n জাতীয়, তাই আরেকভাবেও করা যাবে। তার জন্য প্রথমে দুটো basis-কেই একেকটা matrix-এ পরিণত করে নিতে হবে, basis-এর vector-গুলো column হিসেবে পাশাপাশি লিখে। তবে $\mathcal{B}_1$ থেকে যে matrix-টা পাবে, ধরো তার নাম দিলাম B_1, মানে--

$$B_1 = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 10 \end{bmatrix}.$$

লক্ষ্য করো এটা একটা nonsingular matrix.
একইভাবে $\mathcal{B}_2$-এর জন্যও একটা nonsingular matrix পাবো--

$$B_2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

তবে change of basis matrix-টা হবে $B_2^{-1}B_1$. ■

Exercise 25: আগের অংকে যদি $\mathcal{B}_2$ থেকে $\mathcal{B}_1$-এ যাবার change of basis matrix-টা বার করতে বলতাম, তবে change of basis matrix-টাকে B_1 আর B_2 দিয়ে কীভাবে লিখতে? ■

এই কায়দায় কয়েকটা change of basis matrix বার করে গা গরম করে নেওয়া যাক। তবে কিনা বড় বড় matrix-এর inverse বার করতে হাত ব্যথা হয়ে যাবে, তাই 2×2 সাইজ নিয়ে কাজ করব।

Example 14: $\mathbb{R}^2$-এর দুটো basis দিলাম--

$$\mathcal{B}_1 = \{(1,2),(3,4)\} \text{ এবং } \mathcal{B}_2 = \{(0,-1),(-1,-1)\}.$$

প্রথমটা থেকে দ্বিতীয়টায় যাবার change of basis matrix বার করো। প্রথম basis-এর সাপেক্ষে যে vector হল $(2,3)$ সেটা দ্বিতীয় basis-এর সাপেক্ষে কত?
SOLUTION: এখানে $B_1 = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ আর $B_2 = \begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix}$. আমরা 2×2 সাইজের matrix-দের inverse বার করার একটা ফর্মুলা জানি--

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

সেটা লাগিয়ে $B_2^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}$. সুতরাং change of basis matrix-টা হচ্ছে--

$$B_2^{-1}B_1 = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -1 & -3 \end{bmatrix}.$$

$\mathcal{B}_1$-এর সাপেক্ষে যে vector হল $(2,3)$, সেটা $\mathcal{B}_2$-র সাপেক্ষে

$$\begin{bmatrix} -1 & -1 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -5 \\ -11 \end{bmatrix}.$$

মিলিয়ে দেখে নিতে পারো যে, ঠিক উত্তরই এসেছে--

> $\mathcal{B}_1$-এর সাপেক্ষে $(2,3)$ মানে হল $2 \times (1,2) + 3 \times (3,4) = (11,16)$.
> আবার $\mathcal{B}_2$-র সাপেক্ষে $(-5,-11)$ মানে হল $-5 \times (0,-1) - 11 \times (-1,-1) = (11,16)$.
> দেখলে, দুই ক্ষেত্রেই একই জিনিস এল?

■

Exercise 26: আবার আগের অংকের basis দুটো নিয়েই কাজ করব। কিন্তু এবার $\mathcal{B}_2$ থেকে $\mathcal{B}_1$-এ ফেরার জন্য change of basis matrix-টা বার করো। এবং সেটা ব্যবহার করে দেখো $(-5,-11)$ থেকে সত্যিই $(2,3)$-এ ফিরতে পারছ কিনা।
■

প্রথম ধারণাটা বলা হল। এবার দ্বিতীয় ধারণাটা বলি। Change of basis matrix-দের চেহারা হল $B_2^{-1}B_1$-এর মত, তাই এদের বার করতে একটা matrix-এর inverse করতে হয়, এবং একটা matrix-এর গুণ করতে হয়। যদি matrix-গুলো বড় বড় হয়, তবে কাজটা হাতে বার করা শক্ত। তখন যে সেই সর্বরোগের ধন্বন্তরী RREF-এর ডাক পড়ে, সেটা হয়তো অনুমান করেছ। RREF ব্যবহার করে কাজটা খুবই সহজ। প্রথমে B_2-এর ডানদিকে B_1 লিখে $\begin{bmatrix} B_2 & B_1 \end{bmatrix}$ বানিয়ে নিতে হয়। তারপর এর RREF বার করলে বাঁদিকটা I হয়ে যেতে বাধ্য, মানে RREF-টা দেখতে হবে $\begin{bmatrix} I & P \end{bmatrix}$-এর মত। এই P-টাই হবে change of basis matrix-টা, মানে $B_2^{-1}B_1$. একটা উদাহরণ দেখাই।

Example 15: $\mathbb{R}^3$-এর এই দুটো basis দিচ্ছি--

$$\{(1,2,3),(4,5,6),(7,8,10)\} \text{ এবং } \{(0,0,1),(0,1,1),(1,1,1)\}.$$

প্রথমটা থেকে দ্বিতীয়টাতে যাওয়ার change of basis matrix বার করো।
SOLUTION:

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 4 & 7 \\ 0 & 1 & 1 & 2 & 5 & 8 \\ 1 & 1 & 1 & 3 & 6 & 10 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 4 & 7 \end{array}\right].$$

সুতরাং change of basis matrix-টা হল ওই ডানদিকের অংশটা--

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 4 & 7 \end{bmatrix}.$$

■

এই গেল আমাদের দ্বিতীয় ধারণাটার আলোচনা। এবার তৃতীয় ধারণাটা বলি। এতক্ষণ আমরা খালি F^n জাতীয় vector space-দের নিয়ে কাজ করছিলাম। এবার দেখব যে, RREF-এর কায়দাটা যেকোনো finite dimensional vector space-এর জন্যই কাজ করে।

Example 16: ধরো আমাদের vector space-টা হল সব $a+bx+cx^2$ জাতীয় polynomial-এর, যেখানে $a, b, c \in \mathbb{R}$.

এখানে field-টা $\mathbb{R}$. এর দুটো basis হল $\{1+x, 2, x+x^2\}$ এবং $\{x^2-1, x+3, 4\}$. প্রথমটা থেকে দ্বিতীয়টায় যাওয়ার change of basis matrix বার করো।
SOLUTION: এই অংকটা আমরা একটু আগে গায়ের জোরে করেছি। এবার RREF দিয়ে করব। প্রথম কাজ হল যা খুশি একটা basis নেওয়া, ধরো $\{1, x, x^2\}$. তবে এর সাপেক্ষে $a+bx+cx^2$ হয়ে যাবে (a, b, c). অর্থাৎ এই basis-এর সাপেক্ষে ভাবলে আমাদের vector space-টা আসলে $\mathbb{R}^3$ ছাড়া কিছুই নয়। যে দুটো basis দিয়েছে, তারা হয়ে যাচ্ছে $\mathbb{R}^3$-এর এই দুটো basis--

$$\{(1,1,0), (2,0,0), (0,1,1)\} \text{ এবং } \{(-1,0,1), (3,1,0), (4,0,0)\}.$$

এবার RREF-এর কায়দা লাগালে পাচ্ছি--

$$\left[\begin{array}{ccc|ccc} -1 & 3 & 4 & 1 & 2 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{array}\right].$$

খাড়া লাইনটার ডানদিকের অংশটাই হল change of basis matrix-টা। আগে যখন গায়ের জোরে করেছিলাম, তখনও একই উত্তর পেয়েছিলাম। ■

25.3 Change of basis matrix-দের ধর্ম

একটা change of basis matrix হাতে-কলমে বার করাটা অনেক সময়ে দরকার হয় বটে, কিন্তু তার চেয়েও বেশী দরকার হয়, এদের সরাসরি বার না করে, কেবল বিভিন্ন ধর্ম ব্যবহার করে নাড়াচাড়া করা। ঠিক যেমন যদি একটা nonsingular matrix দিই A, আর বলি AA^{-1} বার করতে, তবে তুমি নিশ্চয়ই A^{-1} বার করতে বসবে না, ধাঁ করে বলে দেবে যে, উত্তর হবে identity matrix. এবার আমরা change of basis matrix-দের এরকম কিছু সুবিধাজনক ধর্ম আলোচনা করব।
একটা basis থেকে অন্য basis-এ যেতে যে change of basis matrix লাগে, ফিরে আসতে লাগে তার inverse-টা[1]। সেটাই নীচের theorem-টার বক্তব্য--

> THEOREM
>
> Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases of a finite dimensional vector space. If the change of basis matrix for going from $\mathcal{B}_1$ to $\mathcal{B}_2$ is P, then the change of basis matrix for going from $\mathcal{B}_2$ to $\mathcal{B}_1$ is P^{-1}.

যদি একটা basis থেকে আরেকটাতে যাও, এবং সেখান ফের অন্য একটা basis-এ যাও, তবে দুই ধাপের change of basis matrix-গুলো গুণ হয়ে যাবে। এটা নিয়েই নীচের theorem-টা--

[1]Inverse বার করতে কোনো অসুবিধা নেই, কারণ change of basis matrix-রা অবশ্যই nonsingular হয়।

> THEOREM
> Let $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$ be three bases of a finite dimensional vector space. If the change of basis matrix from $\mathcal{B}_1$ to $\mathcal{B}_2$ is P_{12}, and the change of basis matrix from $\mathcal{B}_2$ to $\mathcal{B}_3$ is P_{23}, then the change of basis matrix from $\mathcal{B}_1$ to $\mathcal{B}_3$ is $P_{23}P_{12}$.

দুইয়ের বেশী ধাপ থাকলেও একই যুক্তি কাজ করত। খালি এখানে matrix-এর গুণটাতে কে আগে, কে পরে সেটা খেয়াল করো। যে কাজটা আগে করা হচ্ছে, তার matrix-টা ডানদিকে থাকছে।
Change of basis matrix-এর অন্যতম গুরুত্বপূর্ণ ধর্ম হল যে, এরা nonsingular হবেই। এর উল্টো দিকটাও কিন্তু ঠিক, মানে যেকোনো nonsingular matrix-ই একটা change of basis matrix. বিশদ করে বলি--

> ধরো যা খুশি একটা basis আছে, $\mathcal{B}$, আর যা খুশি একটা nonsingular matrix আছে, P. তবে এদের ভিত্তিতে তুমি এমন একটা basis পাবেই (ধরো $\mathcal{D}$), যাতে P হবে $\mathcal{D}$ থেকে $\mathcal{B}$ যাওয়ার change of basis matrix.

একটা উদাহরণ দেখাই।

Example 17: ধরো আমরা $a+bx$ জাতীয় সব polynomial-এর vector space নিয়ে কাজ করছি, যেখানে $a, b \in \mathbb{R}$. এর একটা basis দিচ্ছি $\mathcal{B} = \{1+x, 2+3x\}$, আর একটা nonsingular matrix দিচ্ছি $P = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$. এমন একটা basis দাও, $\mathcal{D}$, যাতে $\mathcal{D}$ থেকে $\mathcal{B}$-তে যাওয়ার change of basis matrix হয় P.
SOLUTION: এখানে $\mathcal{D}$ বানানোর কায়দাটা হল P-এর প্রতিটা column থেকে একটা করে vector তৈরী করা। যেমন--

- P-র প্রথম column-টা $\mathcal{B}$-এর সাপেক্ষে এই vector-টা দেয়--

$$1 \times (1+x) + 2 \times (2+3x) = 5 + 7x.$$

- দ্বিতীয় column-টা দেয়--

$$3 \times (1+x) + 4 \times (2+3x) = 11 + 15x.$$

তাই নতুন basis-টা হবে $\mathcal{D} = \{5+7x, 11+15x\}$.
চাইলে পরীক্ষা করে দেখতে পারো যে, এই $\mathcal{D}$ থেকে $\mathcal{B}$-তে যেতে change of basis matrix-টা P-ই হবে। ■

Exercise 27: আগের অংকে $\mathcal{B} = \{1, 2-x\}$ আর $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ হলে $\mathcal{D}$ কী হত? ■

এই অংক দুটো যে প্রয়োগের দিক থেকে ভীষণ গুরুত্বপূর্ণ, এমন নয়। মোদ্দা যে কথাটা মনে রাখার, সেটা হল--

> THEOREM
> A matrix is a change of basis matrix if and only if it is nonsingular.

25.4 Linear transformation-এর উপর প্রভাব

ধরো দুটো vector space আছে, V আর W, দুজনেই একই field-এর উপরে (ধরো F)। বলে দেওয়া আছে যে, $\dim(V) = n$ আর $\dim(W) = m$. ওদের মধ্যে একটা linear transformation আছে $T : V \to W$, যেটা কোনো $\vec{x} \in V$-কে $T(\vec{x}) \in W$-তে নিয়ে ফেলে (Fig 22)। যদি V-এর একটা basis নিই, $\mathcal{B}_1$ আর W-এর একটা basis নিই $\mathcal{D}_1$, তবে এদের সাপেক্ষে linear transformation-টাকে আমরা একটা matrix দিয়ে প্রকাশ করতে পারি, যাকে Fig 23-এ নাম দিয়েছি A. ছবিটা ভালো করে বুঝে নাও--

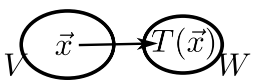


Fig 22

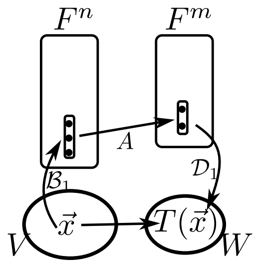


Fig 23

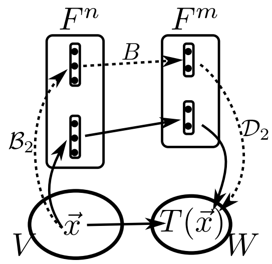


Fig 24

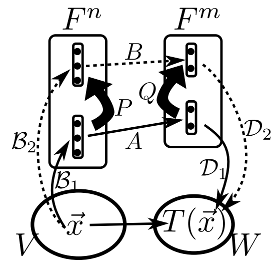


Fig 25

- আমরা প্রথমে $\mathcal{B}_1$ ব্যবহার করে $\vec{x}$-কে F^n-এর মধ্যে একটা tuple-এ পরিণত করছি,
- তারপর সেটাকে $A_{m\times n}$ দিয়ে গুণ করে F^m-এর মধ্যে একটা tuple পাচ্ছি,
- এবং সব শেষে এই tuple-টাকে $\mathcal{D}_1$-এর সাপেক্ষে ব্যবহার করে $T(\vec{x})$-এ পৌঁছচ্ছি।

যদি basis দুটো বদলায় (ধরো $\mathcal{B}_1$-এর জায়গায় $\mathcal{B}_2$, আর $\mathcal{D}_1$-এর বদলে $\mathcal{D}_2$), তবে matrix-টাও বদলে যেতে পারে। ধরো এর নাম দিলাম B. এখন আমরা Fig 24-এ দেখানো ড্যাশ্ ড্যাশ্ পথ ধরে কাজ করছি।
এখানে $\mathcal{B}_1$, $\mathcal{B}_2$ দুজনেই V-এর দুটো basis. ধরো $\mathcal{B}_1$ থেকে $\mathcal{B}_2$-তে যাবার change of basis matrix হল P. একইভাবে $\mathcal{D}_1$ থেকে $\mathcal{D}_2$-তে যাবার change of basis matrix হচ্ছে ধরো Q. ছবিতে দেখিয়েছি Fig 25-এ। এবার আমরা B-কে প্রকাশ করব A, P আর Q দিয়ে। তার জন্য Fig 26-এর দিকে চোখ রাখো। B দিয়ে যা করা যাচ্ছিল, আমরা সেটাই করব P, A আর Q দিয়ে। প্রথমে P-এর তীরটা ধরে উল্টো দিকে যাব, মানে P^{-1} লাগাব। তারপর A লাগাব, সব শেষে Q-এর তীর ধরে উঠব। মানে সব মিলিয়ে দাঁড়ালো

$$B = QAP^{-1}.$$

ভীষণ জটিল লাগল কি? ভালো করে পড়ে বুঝে নাও। যদি notation-এর ঠেলায় মাথা গুলিয়ে যেতে চায়, তবে তার প্রতিষেধক হিসেবে নীচের অংকটা করো।

Exercise 28: ধরো domain-এর basis দুটো $\mathcal{A}$ আর $\mathcal{B}$, এবং codomain-এর basis দুটো $\mathcal{D}$ আর $\mathcal{E}$. বলে দিচ্ছি যে, P_{BA} হল $\mathcal{B}$ থেকে $\mathcal{A}$-তে যাবার change of basis matrix, আর P_{ED} হল $\mathcal{E}$ থেকে $\mathcal{D}$-তে যাবার change of basis matrix. একটা linear transformation আছে, $\mathcal{A}, \mathcal{D}$-র সাপেক্ষে যার matrix হল M. তোমাকে বলতে হবে $\mathcal{B}, \mathcal{E}$-র সাপেক্ষে তার matrix-টা কী হবে।
HINT: এটা আসলে আগের আলোচনাটাই, খালি notation বদলে দিয়েছি, যাতে তুমি নতুন করে ভাবতে বাধ্য হও। ছবি এঁকে না নিলে মাথা গুলিয়ে যাবে কিন্তু! ■

Fig 26

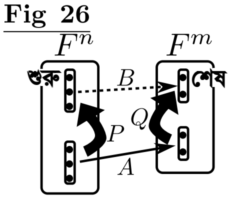
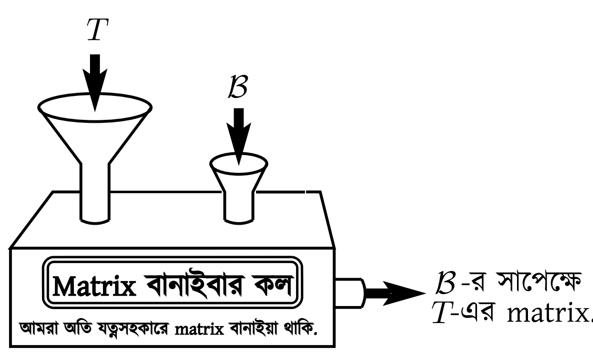


Fig 27



মোদ্দা কথাটা এই যে, basis বদলালে আগের matrix-টার দুদিকে দুটো change of basis matrix দিয়ে গুণ হয়ে যায়। এর উল্টো দিকটাও সত্যি, কারণ আমরা আগেই দেখেছি যে, যেকোনো nonsingular matrix-কেই change of basis matrix বলে ভাবা যায়। তাই যদি A যেকোনো matrix এবং P, Q যেকোনো দুটো nonsingular matrix নাও, তবে PAQ আর A একই linear transformation-কে বোঝাবে, খালি হয়তো basis দুটোকে পরিবর্তন করে নিতে হবে।
এবার একটা বিশেষ ক্ষেত্র আলোচনা করব, যেটা অত্যন্ত গুরুত্বপূর্ণ। সেটা হল যখন domain আর codomain একই, এবং দুই দিকেই আমরা একই basis ব্যবহার করছি। এই বিশেষ ক্ষেত্রটা বারবার কাজে লাগবে। এখান থেকেই linear algebra-র একটা ভীষণ গুরুত্বপূর্ণ ধারণার জন্ম, যার নাম **similarity.**

25.5 Similarity

ধরো V হল একটা finite dimensional vector space, আর $T : V \to V$ হল একটা linear transformation. তোমাকে V-এর দুটো basis দেওয়া আছে, $\mathcal{B}_1$ আর $\mathcal{B}_2$. ধরো P হল $\mathcal{B}_1$ থেকে $\mathcal{B}_2$-তে যাবার change of basis matrix. যদি $\mathcal{B}_1$-এর সাপেক্ষে T-এর matrix হয় A, তবে $\mathcal{B}_2$-র সাপেক্ষে হবে PAP^{-1}. এর উল্টো দিকটাও সত্যি, অর্থাৎ যদি P যেকোনো nonsingular matrix হয়, তবে PAP^{-1} ওই একই linear transformation-কেই প্রকাশ করবে (অন্য কোনো basis-এর সাপেক্ষে)।
এই দুদিক মিলিয়ে পাব এই অত্যন্ত গুরুত্বপূর্ণ theorem-টা--

> **THEOREM**
>
> Two square matrices A, B both represent the same linear transformation (w.r.t. two different bases) if and only if
> $$B = PAP^{-1}$$
> for some nonsingular matrix P.

এর সঙ্গেই ওতপ্রোতভাবে মিলে আছে এই সংজ্ঞাটা--

> **DEFINITION: Similar**
>
> Two matrices A and B are called **similar**, if $B = PAP^{-1}$ for some nonsingular matrix P.

এই সংজ্ঞাটা ব্যবহার করে আগের theorem-টাকে বলা যায় এইভাবে--

> Two matrices both represent the same linear transformation (wrt possibly different bases) if and only if they are similar.

Basis বদলালে matrix-টা কীভাবে বদলায়, সেটা বার করতে আমাদের আগ্রহ কেন, সেটা আগে বলে নিই। T-কে একটা matrix দিয়ে প্রকাশ করার জন্য V-এর একটা কোনো basis নিতে হচ্ছে, $\mathcal{B}$. ব্যাপারটাকে Fig 27-এর মত করে ভাবতে পারো। এই matrix-টার মধ্যে দুরকম জিনিস মিশে আছে-- linear transformation-টা এবং basis-টা। এদের মধ্যে linear transformation-টা নিয়েই আমাদের আগ্রহ বেশী, basis-টা তো খালি একটা coordinate system, লেখার সুবিধার জন্যই খালি ওর অবতারণা। সুতরাং আমরা matrix-টার সেই সব ধর্ম নিয়ে বেশী আগ্রহী হব, যেগুলো linear transformation-টা থেকে আসে, basis-টা থেকে নয়। অর্থাৎ সেই সব ধর্ম যেগুলো basis বদলালেও অপরিবর্তিত থাকে। আমরা যখন matrix-এর বিভিন্ন ধারণা নিয়ে আলোচনা করি, তখন আমাদের আসল উদ্দেশ্য থাকে সেই matrix-টা যে linear transformation-টার প্রতিনিধি, সেটার ধর্মগুলো নিয়ে আলোচনা করা। নইলে খামোখা একগুচ্ছ সংখ্যাকে row আর

column-এ সাজিয়ে লিখে matrix বানিয়ে তার ধর্মাধর্ম নিয়ে মাথা ঘামাব কেন? এই প্রশ্নটা হয়তো প্রথম যখন হায়ার সেকণ্ডারীর সময়ে matrix শিখেছিলে তখনই মনে এসেছিল। আমি যখন প্রথম trace-এর সংজ্ঞা শিখেছিলাম, তখন অবাক হয়ে ভেবেছিলাম যে, খালি খালি এই diagonal বরাবর entry-গুলোর যোগফল নিয়ে এত মাতামাতি করার কী আছে বাপু? কই অন্য diagonal-টার entry-গুলোকে কখনো যোগ করছি না! এসব প্রশ্নেরই উত্তর রয়েছে similar matrix-এর ধারণার মধ্যে। যদি দুটো similar matrix নাও তবে তাদের trace সমান হতে বাধ্য। তার মানে trace হল আসলে linear transformation-এর ধর্ম, basis বদলালেও বদলে যায় না। কিন্তু অন্য diagonal-টার entry-র যোগফল basis বদলালেই বদলে যেতে পারে। তাই ওটা নিয়ে আমাদের আগ্রহ নেই। একই ব্যাপার ঘটে rank-এর বেলাতেও। এরা সকলেই আসলে linear transformation-এর ধর্ম।

DAY 26
Linear transformation-এর অঙ্গ-প্রত্যঙ্গ

যেকোনো linear transformation-এর সঙ্গে কয়েকটা গুরুত্বপূর্ণ ধারণা জড়িত থাকে, যেগুলোর সম্বন্ধে পরিষ্কার ধারণা থাকা আবশ্যক।

26.1 Image

Fig 28-এর দিকে তাকাও। ধরো $f: V \to W$ হল একটা linear transformation. তার মানে প্রথম কথা হল f হচ্ছে V থেকে W-তে একটা function. এমন বলিনি যে f-কে onto হতে হবে, তাই W-এর মধ্যে সব value-ই যে f নেবে, এমন কোনো কথা নেই। যে সব value নেবে, তাদের set-কে বলে f-এর **image.** ছবিতে এটাকে W-এর পেটের মধ্যে শেড করে দেখিয়েছি। যদি f-টা onto হত, তবে এই image-টা পুরো W জুড়ে থাকত, মানে পুরো W-টাই শেড হয়ে যেত।

DEFINITION: **Image**

Let $T: V \to W$ be a linear transformation. Then its **image** is the defined as

$$\mathrm{Im}(T) = \{T(\vec{v}) \ : \ \vec{v} \in V\}.$$

Example 18: If $T: \mathbb{R}^3 \to \mathbb{R}^2$ is given by

$$T(x, y, z) = (x + y - z, 2x + 2y - 2z),$$

then find $\mathrm{Im}(T)$.

Fig 28

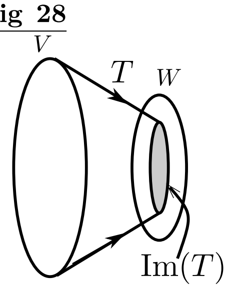
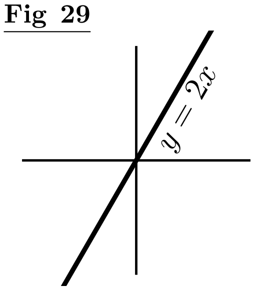


Fig 29



SOLUTION: প্রথমে কয়েকটা উদাহরণ নিয়ে আন্দাজ করে নিই $\text{Im}(T)$ দেখতে কীরকম। ধরো $(x,y,z)=(1,2,3)$ নিলাম। তবে $T(x,y,z)=(0,0)$ হবে। যদি $(x,y,z)=(1,1,1)$ হয়, তবে $T(x,y,z)=(1,2)$. লক্ষ্য করো সর্বদাই প্রথম সংখ্যাটার দ্বিগুণ হচ্ছে দ্বিতীয় সংখ্যাটা। সর্বদাই এটা হবে, কারণ $2x+2y-2z=2(x+y-z)$. প্রথম সংখ্যাটা যেকোনো real number হতে পারে। তাই--

$$\text{Im}(T)=\{(x,y)\ :\ y=2x\}.$$

ছবিতে দেখিয়েছি Fig 29-এ। ■

Exercise 29: If $T:\mathbb{R}^3\to\mathbb{R}^3$ is given by

$$T(x,y,z)=(x+y,y+z,x-z),$$

then find $\text{Im}(T)$. Check if $(1,2,3)$ and $(1,2,-1)$ are in $\text{Im}(T)$. ■

আমরা এ যাবৎ এ কথাটা ব্যবহার করিনি যে, T একটা linear transformation. যেকোনো function-এর জন্যই image একইভাবে define করা হয়। কিন্তু T-টা linear transformation হওয়ার দরুণ $\text{Im}(T)$-এর একটা বাড়তি বৈশিষ্ট্য থাকবে, এটা একটা subspace হবে। সেটাই বলেছে নীচের theorem-টায়।

THEOREM

Let V and W be two vector spaces over a field F. Let $T:V\to W$ be a linear transformation. Then $\text{Im}(T)$ is a subspace of W.

Proof: কোনো subset-কে subspace দেখানোর জন্য কী করতে হয়, সে তো জানোই। প্রথমে দেখাতে হয় nonempty. তারপর যেকোনো $\vec{u},\vec{v}\in\text{Im}(T)$ নিতে হয়, যেকোনো $\alpha,\beta\in F$ নিতে হয়, এবং দেখাতে হয় যে, $\alpha\vec{u}+\beta\vec{v}\in\text{Im}(T)$ হতে বাধ্য।

Here $\text{Im}(T)\neq\phi$, since $T(\vec{0}_V)=\vec{0}_W\in\text{Im}(T)$.

So enough to show that

$$\forall\alpha,\beta\in F\quad\forall\vec{u},\vec{v}\in\text{Im}(T)\quad\alpha\vec{u}+\beta\vec{v}\in\text{Im}(T).$$

$\forall\alpha,\beta$ Take any $\alpha,\beta\in F$.

$\forall\vec{u},\vec{v}$ Take any $\vec{u},\vec{v}\in\text{Im}(T)$.

Then

$$\exists\vec{x},\vec{y}\in V\quad T(\vec{x})=\vec{u},\quad T(\vec{y})=\vec{v}.$$

এটা হল কারণ $\vec{u},\vec{v}$ রয়েছে $\text{Im}(T)$-র মধ্যে, তার মানে T এই value দুটো নেয়, অর্থাৎ কোনো $\vec{x},\vec{y}\in V$-এর জন্য $T(\vec{x})=\vec{u}$ আর $T(\vec{y})=\vec{v}$.

So $\alpha\vec{x}+\beta\vec{y} \in V$, and

$$\begin{aligned} T(\alpha\vec{x}+\beta\vec{y}) &= \alpha T(\vec{x})+\beta(\vec{y}) \quad [\because T \text{ is linear}] \\ &= \alpha\vec{u}+\beta\vec{v}. \end{aligned}$$

Thus $\alpha\vec{u}+\beta\vec{v} \in \mathrm{Im}(T)$, as required.

[Q.E.D]

যেহেতু $\mathrm{Im}(T)$ একটা subspace, তাই তার একটা dimension আছে। এর একটা বিশেষ নাম আছে--

DEFINITION: **Rank**

If $T : V \to W$ is a linear transformation, then $\dim(\mathrm{Im}(T))$ is called the **rank** of T.

Example 19: একটু আগে আমরা একটা অংকে $T(x, y, z) = (x + y - z, 2x2y - 2z)$ নিয়ে কাজ করেছিলাম। সেখানে দেখেছিলাম

$$\mathrm{Im}(T) = \{(x, y) \ : \ y = 2x\}.$$

এখানে T-এর rank কত?

SOLUTION: এখানে $\mathrm{Im}(T) = \mathrm{span}\{(1, 2)\}$, তাই $\dim(\mathrm{Im}(T)) = 1$. সুতরাং rank হল 1. ■

Exercise 30: যদি $T : \mathbb{R}^3 \to \mathbb{R}^3$ হয় $T(x, y, z) = (x + y, y + z, x + z)$, তবে T-এর rank কত হবে? ■

Example 20: If $T : \mathbb{R}^4 \to \mathbb{R}^3$ is defined as $T(x, y, z, t) = (x-y-z-t, x+2z-t, x+y+3z-3t)$, then find the rank of T.[3] **(2021 (part))**

SOLUTION:

Every vector in $\mathrm{Im}(T)$ is like

$$\begin{aligned} T(x, y, z, t) &= (x-y-z-t, x+2z-t, x+y+3z-3t) \\ &= x(1, 1, 1) + y(-1, 0, 1) + z(-1, 2, 3) + t(-1, -1, -3). \end{aligned}$$

So $\mathrm{Im}(T) = \mathrm{span}\{(1, 1, 1),\ (-1, 0, 1),\ (-1, 2, 3),\ (-1, -1, -3)\}$.

Now $(1, 1, 1)$ and $(-1, 0, 1)$ are not multiples of each other. So $\dim(\mathrm{Im}(T)) \geq 2$.

Let, if possible, $(-1, 2, 3) = a(1, 1, 1) + b(-1, 0, 1)$ for some $a, b \in \mathbb{R}$. Then $a - b = -1,\ a = 2$ and $a + b = 3$.

From the first two equations, $a = 2$ and $b = 3$. But these do not satisfy the third equation.

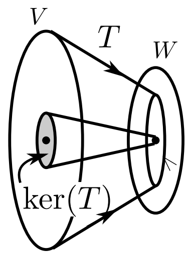


Fig 30

> So $\dim(\mathrm{Im}(T)) \geq 3$. But $\mathrm{Im}(T) \subseteq \mathbb{R}^3$. Hence $\dim(\mathrm{Im}(T)) = 3$.
> Hence $\mathrm{rank}(T) = 3$.

■

26.2 Kernel

আমরা কাজ করছি একটা linear transformation নিয়ে, $T : V \to W$. মনে রেখো যে V আর W কেবল দুটো set-ই নয়, এরা vector space-ও বটে। সুতরাং দুজনের মধ্যেই zero vector আছে। এদেরকে Fig 30-এ দুটো কালো ডট্ দিয়ে দেখিয়েছি। লেখার সময়ে $\vec{0}_V$ আর $\vec{0}_W$ লিখব। আমরা জানি যে f-টা linear transformation বলে $f(\vec{0}_V) = \vec{0}_W$ হবে। অবশ্য f-টা তো one-one বলা নেই, তাই অন্য কোনো $\vec{v} \in V$-ও পাওয়া যেতে পারে যাতে $f(\vec{v}) = \vec{0}_W$ হয়। এরকম যাবতীয় $\vec{v}$-এর set-এর একটা নাম আছে--

> DEFINITION: **Kernel/nullspace**
>
> The **kernel** or **nullspace** of a linear transformation $T : V \to W$ is defined as
>
> $$\ker(T) = \{\vec{v} \in V \ \ : \ \ T(\vec{v}) = \vec{0}_W\}.$$

Fig 30-এ শেড করে দেখিয়েছি। লক্ষ্য করো, $\mathrm{Im}(T)$ ছিল W-এর মধ্যে, আর $\ker(T)$ আছে V-এর মধ্যে।

Example 21: একটা linear transformation দিচ্ছি, $T : \mathbb{R}^3 \to \mathbb{R}^2$, যেখানে

$$T(x_1, x_2, x_3) = (x_1 + x_2, x_3).$$

তবে $\ker(T)$ কী হবে?

SOLUTION: $(x_1, x_2, x_3) \in \ker(T)$ হওয়া মানে $T(x_1, x_2, x_3) = (0, 0)$ হওয়া, অর্থাৎ $x_1 + x_2 = 0$ এবং $x_3 = 0$ হওয়া, অর্থাৎ কিনা x_1 যা খুশি হতে পারে, কিন্তু $x_2 = -x_1$ আর $x_3 = 0$ হলেই হল। তার মানে $(x_1, -x_1, 0)$-র মত দেখতে যারা, খালি তাদের নিয়েই $\ker(T)$ তৈরী, অর্থাৎ--

$$\ker(T) = \{(x, -x, 0) \ \ : \ \ x \in \mathbb{R}\}.$$

■

Example 22: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear mapping defined by $T(x, y, z) = (x+y, y+z, z+x)$, $(x, y, z) \in \mathbb{R}^3$. Then kernel of T is equal to

(A) $\{(1, 0, 0)\}$

(B) $\{(1, 0, 0), (0, 1, 0)\}$

(C) $\{\theta\}$ (θ being the null vector $(0, 0, 0)$)

(D) none of the above.

(2019.CC6.1h)
SOLUTION:প্রথমে একটা জিনিস আমাকে বলো--যদি T যা খুশি একটা linear transformation হয়, এবং $T(\vec{v}) = \vec{0}$ হয়, তবে $T(2\vec{v})$ কত হবে? নিশ্চয়ই সন্দেহ নেই যে, $T(2\vec{v}) = 2T(\vec{v}) = 2\vec{0} = \vec{0}$ হবে। সুতরাং যদি $\vec{v} \in \ker(T)$ হয়, তবে $2\vec{v}$-ও $\ker(T)$-এর মধ্যেই থাকতে বাধ্য। সুতরাং প্রথম দুটো option এখানেই বাদ হয়ে গেল।
যাই হোক, এরকম option বাদ দেওয়ার চেষ্টা না করে অংকটা সরাসরি করে ফেলা যথেষ্ট সহজ।

$$(x, y, z) \in \ker(T) \iff (x+y, y+z, z+x) = (0, 0, 0) \iff x = y = z = 0.$$

তাই উত্তর হবে (C). ■

একইরকম একটা অংক তোমার জন্য।

Exercise 31: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear mapping defined by $T(x, y, z) = (x+y, y-z, z+x)$, $(x, y, z) \in \mathbb{R}^3$. Then kernel of T is equal to

(A) $\text{span}\{(1, -1, -1)\}$

(B) $\{(1, -1, -1), (0, -1, 1)\}$

(C) $\{\vec{0}\}$

(D) $\text{span}\{(0, -1, 1)\}$.

■

Exercise 32: একটা linear transformation-এর kernel-এ খালি একটাই vector আছে। তবে বলতে পারো kernel-টা কী? ■

এবার একটু সামান্য অন্যরকম অংক। একটা linear transformation দেওয়া থাকলে তার kernel বার করা শিখেছ। এবার kernel-টা দিয়ে দেবে, তোমাকে সেই kernel-ওয়ালা একটা কোনো linear transformation বার করতে হবে।

Example 23: Find a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that the set of all vectors (x_1, x_2, x_3) satisfying the equation $4x_1 - 3x_2 + x_3 = 0$ is $\ker(T)$.[5] **(2021.CC12.3b)**
SOLUTION: এই অংকটা আসলে ভীষণই সহজ, চোখে দেখেই ধাঁ করে উত্তর লিখে দেওয়া যায়। প্রথমে ধরো $T : \mathbb{R}^3 \to \mathbb{R}$ (মানে codomain-টাকে $\mathbb{R}^3$-এর বদলে $\mathbb{R}$ নিলাম)। তবে $T(x_1, x_2, x_3)$-এর চেহারা হবে $ax_1 + bx_2 + cx_3$ জাতীয়।

প্রশ্ন হল--এমন কী a, b, c আছে, যাতে $4x_1 - 3x_2 + x_3 = 0$ হওয়া আর $ax_1 + bx_2 + cx_3 = 0$ হওয়া একই জিনিস হয়? উত্তর তো বলাই বাহুল্য, $a = 4,\ b = -3$ আর $c = 1$ নিলেই হয়! বা তার কোনো nonzero multiple, যেমন $a = 8,\ b = -6$ আর $c = 2$. এবার ধরো বললাম $T : \mathbb{R}^3 \to \mathbb{R}^2$ (মানে codomain-টাকে $\mathbb{R}^2$ করেছি)। এবার $T(x_1, x_2, x_3)$-এর মধ্যে দুটো component, দুজনেই সেই $ax_1 + bx_2 + cx_3$ জাতীয়। সুতরাং দুজনের ক্ষেত্রেই আগের যুক্তি খাটবে। আমাদের অংকে codomain-টা $\mathbb{R}^3$ দিয়েছে, মানে তিনটে component. সবার বেলাতেই একই কায়দা লাগানো যাবে, তাই--

> One such linear transformation is
>
> $$T(x_1, x_2, x_3) = \begin{bmatrix} 4 & -3 & 1 \\ 4 & -3 & 1 \\ 4 & -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

এটাই অবশ্য একমাত্র উত্তর নয়। Matrix-টার row-গুলোকে $(4, -3, 1)$-এর যা খুশি nonzero multiple নেওয়া যায়। অংকটায় যদিও কিছু প্রমাণ করতে বলেনি, তাও প্রমাণটা করে দেওয়াই উচিত হবে।

> Shall show $\ker(T) = \{(x_1, x_2, x_3) \in \mathbb{R}^3\ :\ 4x_1 - 3x_2 + x_3 = 0\} = S$, say.

দুটো set-কে সমান দেখাতে হলে দেখাতে হয় যে, এই set-এর যেকোনো element ওই set-এ আছে, আবার ওই set-এর যেকোনো element আছে এই set-এ।

> Take any $(x_1, x_2, x_3) \in \ker(T)$.
>
> Then $T(x_1, x_2, x_3) = (0, 0, 0)$,
>
> ie, $\begin{bmatrix} 4 & -3 & 1 \\ 4 & -3 & 1 \\ 4 & -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.
>
> Hence $4x_1 - 3x_2 + x_3 = 0$.
>
> So $(x_1, x_2, x_3) \in S$.
>
> Hence $\ker(T) \subseteq S$.

এবার উল্টোদিকটা--

> Next, take any $(x_1, x_2, x_3) \in S$.
>
> Then $4x_1 - 3x_2 + x_3 = 0$.
>
> So $\begin{bmatrix} 4 & -3 & 1 \\ 4 & -3 & 1 \\ 4 & -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$,
>
> or $T(x_1, x_2, x_3) = (0, 0, 0)$.
>
> Hence $(x_1, x_2, x_3) \in \ker(T)$.
>
> Thus, $S \subseteq \ker(T)$.

দু দিক মিলিয়ে পাচ্ছি--

Combining, we have $\ker(T) = S$, as required.

■

Image-এর মত kernel-ও একটা subspace হয়। সেটা নিয়েই নীচের theorem-টা।

> **THEOREM**
>
> Let V and W be two vector spaces over a field F. Let $T : V \to W$ be a linear transformation. Then $\ker(T)$ is a subspace of W.

Proof:

By definition, $\ker(T) \subseteq V$. Also $\ker(T) \neq \phi$, since $T(\vec{0}_V) = \vec{0}_W$, and hence $\vec{0}_V \in \ker(T)$.

Enough to show that

$\forall \alpha, \beta \in F \quad \forall \vec{u}, \vec{v} \in \ker(T) \quad \alpha\vec{u} + \beta\vec{v} \in \ker(T)$.

$\forall \alpha, \beta$ Take any $\alpha, \beta \in F$.

$\forall \vec{u}, \vec{v}$ Take any $\vec{u}, \vec{v} \in \ker(T)$.

Then

$$\begin{aligned} T(\alpha\vec{u} + \beta\vec{v}) &= \alpha T(\vec{u}) + \beta T(\vec{v}) \quad [\because T \text{ is linear}] \\ &= \alpha\vec{0}_W + \beta\vec{0}_W \quad [\because \vec{u}, \vec{v} \in \ker(T)] \\ &= \vec{0}_W. \end{aligned}$$

Thus $\alpha\vec{u} + \beta\vec{v} \in \ker(T)$, as required.

[Q.E.D]

Example 24: Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ which maps the basis vectors $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ of $\mathbb{R}^3$ to the vectors $(2, 0, 0)$, $(0, 2, 0)$, $(0, 0, 2)$ respectively. Find $\ker(T)$ and $\text{Im}(T)$.[2+1+1] **(2019(part))**

SOLUTION:প্রথমে T-এর ফর্মুলা বার করতে বলেছে, মানে যেকোনো (x, y, z)-এর উপরে T লাগালে কী হবে, সেটা বার করতে হবে। তার জন্য (x, y, z)-কে ওই basis-এর সাপেক্ষে লিখে নিতে হবে--

Let $(x, y, z) = a(0, 1, 1) + b(1, 0, 1) + c(1, 1, 0)$.

খানিকক্ষণ লড়াই করার পর পাবে--

> Then $a = \frac{1}{2}(-x+y+z)$, $b = \frac{1}{2}(x-y+z)$ and $c = \frac{1}{2}(x+y-z)$.
> So $T(x,y,z) = aT(0,1,1) + bT(1,0,1) + cT(1,1,0) = \cdots = 2(a,b,c) = (-x+y+z, x-y+z, x+y-z)$.

এবার kernel বার করব--

> $T(x,y,z) = (0,0,0) \iff (-x+y+z, x-y+z, x+y-z) = (0,0,0) \iff x = y = z = 0$.
> Hence $\ker(T) = \{(0,0,0)\}$.

এবার image-এর পালা--

$$\begin{aligned} \mathrm{Im}(T) &= \{(-x+y+z, x-y+z, x+y-z) \ : \ x,y,z \in \mathbb{R}\} \\ &= \{x(-1,1,1) + y(1,-1,1) + z(1,1,-1) \ : \ x,y,z \in \mathbb{R}\} \\ &= \mathrm{span}\{(-1,1,1),\ (1,-1,1),\ (1,1,-1)\} \\ &= \mathbb{R}^3, \end{aligned}$$

> Since $\{(-1,1,1),\ (1,-1,1),\ (1,1,-1)\}$ is linearly independent.

■

Exercise 33: Determine the linear mapping $T:\mathbb{R}^3 \to \mathbb{R}^3$ that maps the basis vectors $(0,1,1)$, $(1,0,1)$, $(1,1,0)$ of $\mathbb{R}^3$ to the vectors $(2,1,1)$, $(1,2,1)$, $(1,1,2)$ respectively. Find $\ker(T)$ and $\mathrm{Im}(T)$.[2+1+1] **(2020(part))**
HINT: আগের অংকের মত করো। ■

Image-এর dimension-কে যেমন rank বলে, তেমনি kernel-এর dimension-কে বলে একটা linear transformation-এর **nullity.**

> DEFINITION: **Nullity**
>
> The dimension of the kernel of a linear transformation is called the **nullity** of the linear transformation.

Example 25: If $T:\mathbb{R}^4 \to \mathbb{R}^3$ is defined as $T(x,y,z,t) = (x-y-z-t, x+2z-t, x+y+3z-3t)$,

then find the rank and nullity of T.[3+2] **(2021)**
SOLUTION: আমরা rank-টা বার করেছিলাম Example 20-তে[2]। এবার nullity বার করব।

[2] সেটা ছিল 212 নম্বর পাতার ধারেকাছে কোথাও।

$(x, y, z, t) \in \ker(T) \iff (x - y - z - t, x + 2z - t, x + y + 3z - 3t) = (0, 0, 0)$, ie

$$\begin{aligned} x - y - z - t &= 0, \\ x + 2z - t &= 0, \\ x + y + 3z - 3t &= 0. \end{aligned}$$

From the first equation $t = -x + y + z$.
From the second, $x = t - 2z = -x + y + z - 2z = -x + y - z$,
or $x = \frac{1}{2}(y - z)$.
So $t = -x + y + z = \frac{1}{2}(y + 3z)$.
From the third equation, $\frac{1}{2}(y - z) + y + 3z - \frac{3}{2}(y + 3z) = 0$,
or $z = 0$.
So $x = t = \frac{y}{2}$. Thus $\ker(T) = \left\{ y\left(\frac{1}{2}, 1, 0, \frac{1}{2}\right) \ : \ y \in \mathbb{R} \right\} = \text{span}\left\{ \left(\frac{1}{2}, 1, 0, \frac{1}{2}\right) \right\}$.
Clearly, this has dimension 1. So $\text{nullity}(T) = 1$.

এখানে আমরা সরাসরি সংজ্ঞা থেকে nullity বার করে দেখালাম। আসলে কিন্তু অংকটা একটা কৌশল করে অনেক সহজে করে ফেলা যায়। Rank আর nullity-র মধ্যে একটা বিশেষ সম্পর্ক আছে, যা দিয়ে একটা থেকে অন্যটা বার করে ফেলা যায়। সম্পর্কটা আমরা পরশু শিখব। তাতে rank বার হয়ে যাবার পর মাত্র এক লাইনে nullity বার হয়ে যাবে। ■

Example 26: Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2, x_1 - x_2, 0).$$

Then

(A) $\dim(\ker(T)) = 2$

(B) $\dim(\text{Im}(T)) = 2$

(C) $\ker(T) = \text{Im}(T)$

(D) $\ker(T) \subsetneq \text{Im}(T)$

(2022)
SOLUTION: প্রথমে kernel-টা বার করে নেওয়া যাক। $(x_1, x_2, x_3) \in \ker(T) \iff x_1 - x_2 = 0$. তার মানে

$$\ker(T) = \{(x_1, x_2, x_3) \ : \ x_1 = x_2\}.$$

এটা যে Fig 31-এর plane-টা, সেটা আশা করি বুঝতে পারছ। সুতরাং (A) ঠিক হচ্ছে।
এবার image-টা দেখি--

$$\text{Im}(T) = \{(x_1 - x_2, x_1 - x_2, 0) \ : \ x_1, x_2 \in \mathbb{R}\} = \{(x, x, 0) \ : \ x \in \mathbb{R}\}.$$

এই শেষের ধাপটা পেলাম $x = x_1 - x_2$ বসিয়ে। সুতরাং image-টা হল $\text{span}\{(1, 1, 0)\}$. তাই ওর dimension হচ্ছে 1. তাই (B) আর (C) একসঙ্গে কচ্ করে নাকচ হয়ে গেল। সেই সঙ্গে (D)-ও পটল তুলল, কারণ একটা 2 dimension-এর

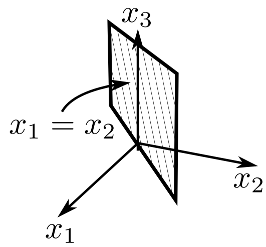


Fig 31

subspace কী করে আর একটা 1 dimension-এর subspace-র ভিতরে থাকে? ■

Exercise 34: Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2, x_1 - x_2, x_3).$$

Then which of the following is/are true?

(A) $\dim(\ker(T)) = 2$

(B) $\dim(\mathrm{Im}(T)) = 2$

(C) $\ker(T) = \mathrm{Im}(T)$

(D) $\ker(T) \subsetneq \mathrm{Im}(T)$

■

Example 27: Let V be a vector space of 2×2 matrices over $\mathbb{R}$. Let $T : V \to V$ be the linear mapping defined by

$$T(A) = AM - MA \text{ where } M = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}.$$

Find a basis of $\ker(T)$ and the dimension of it.

SOLUTION:

$\ker(T) = \{A \in V \ : \ T(A) = O\}$, where is O is 2×2 zero matrix.

Now

$$\begin{aligned} T(A) = O \quad &\iff \quad AM = MA \\ &\iff \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} \\ &\iff \begin{bmatrix} a & 2a+3b \\ c & 2c+3d \end{bmatrix} = \begin{bmatrix} a+2c & b+2d \\ 3c & 3d \end{bmatrix} \\ &\iff \begin{bmatrix} 2c & 2d-2a-2b \\ 2c & -2c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ &\iff c = 0, \ \ d = a + b. \end{aligned}$$

Thus

$$\ker(T) = \left\{ \begin{bmatrix} a & b \\ 0 & a+b \end{bmatrix} \ : \ a, b \in \mathbb{R} \right\}.$$

Now

$$\begin{bmatrix} a & b \\ 0 & a+b \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Also

$$a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \iff a = b = 0.$$

Thus

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \right\}$$

is a linearly independent set generating $\ker(T)$, and hence is a basis of $\ker(T)$.

The dimension is the size of the basis, ie, 2.

■

DAY 27
Matrix-এর সঙ্গে সম্পর্ক

ধরো $T : V \to W$ হল একটা linear transformation. তবে T-এর সঙ্গে জড়িত বিভিন্ন ধারণার কথা আমরা জেনেছি, যেমন $\mathrm{Im}(T)$, $\ker(T)$ এবং এদের dimension-গুলো, মানে rank এবং nullity. যদি V আর W দুজনেই finite dimensional হয় (ধরো $\dim(V) = n$ আর $\dim(W) = m$), তবে V আর W-এর যেকোনো দুটো basis নিলে তাদের সাপেক্ষে T-কে একটা $m \times n$ সাইজের matrix দিয়ে প্রকাশ করা যায়। আন্দাজ করা কঠিন নয় যে, এই matrix-টার বিভিন্ন ধর্মের সঙ্গে linear transformation-টার বিভিন্ন ধর্মের ওতপ্রোত যোগ থাকবে। সেই ব্যাপারটাই আজকে আলোচনা করব।

27.1 Column space বনাম image

কোনো linear transformation-এর image থেকে আসে তার matrix-এর column space. ব্যাপারটা বোঝার জন্য একটা উদাহরণ মাথায় রাখলে সুবিধা হবে।

Example 28: আমরা কাজ করব $T : V \to W$ নিয়ে, যেখানে V আর W দুজনেই হল যাবতীয় $ax^2 + bx + c$ জাতীয় polynomial-এর vector space, যেখানে $a, b, c \in \mathbb{R}$. আমরা linear transformation-টা নেব differentiation. যেমন $T(2x^2 + 5x - 6) = 4x + 5$ হবে। আমরা V-এর (এবং W-এরও) basis নেব $\{1, x, x^2\}$. এই basis-এর সাপেক্ষে T-এর matrix-টা কী হবে?

SOLUTION: যেই কোনো $\vec{x} \in V$ নেব, সেটা এই basis-এর সাপেক্ষে একটা 3-tuple-এর জন্ম দেবে, যেমন $2x^2 + 5x - 6$ পরিণত হবে $(-6, 5, 2)$-এ। সুতরাং এই basis-এর সাপেক্ষে V এবং W দুজনেই যেন $\mathbb{R}^3$.

যেহেতু $T(ax^2 + bx + c) = 2ax + b$, তাই linear transformation-টা যেন (c, b, a)-কে পাঠাচ্ছে $(b, 2a, 0)$-তে। এ থেকেই matrix-টা পেয়ে যাব--

$$\begin{bmatrix} b \\ 2a \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}}_{A} \begin{bmatrix} c \\ b \\ a \end{bmatrix}.$$

■

এই অংকে $\vec{x}$-কে যদি $\mathbb{R}^3$-এর সদস্য হিসেবে ভাবি, তবে $T(\vec{x})$-কে ভাবব $A\vec{x}$ হিসেবে। ফলে $\text{Im}(T)$ তৈরী হবে সেই সব $\vec{w} \in W$ দিয়ে যাদেরকে $\mathbb{R}^3$-এর সদস্য হিসেবে ভাবলে $A\vec{x}$ আকারে লেখা যায়। সেই জন্যই বলতে পারি--

> $\text{Im}(T)$ থেকেই আসছে $\mathscr{C}(A)$.

সাবধান, $\text{Im}(T)$ আর $\mathscr{C}(A)$ সমান বলছি না কিন্তু! হাজার হোক, $\text{Im}(T)$-এর সদস্যরা হল একেকটা polynomial, আর $\mathscr{C}(A)$-র সদস্যরা হল একেকটা 3-tuple! তারা সমান হয় কী করে? যদি basis-টা বদলাও, তবে A-ও বদলাবে, ফলে $\mathscr{C}(A)$-ও বদলে যেতে পারে। কিন্তু প্রতিটা basis-এর জন্যই আলাদা করে আমাদের আলোচনাটা খাটবে।

Exercise 35: ফের একই $T : V \to W$ নিয়ে কাজ করব। কিন্তু এবার basis-টা নেব $\{x^2, x, 1\}$. এবার যে matrix-টা পাবে, সেটা হল

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

দেখাও যে $\mathscr{C}(A) \neq \mathscr{C}(B)$. ■

27.2 Nullspace বনাম kernel

যেরকম যুক্তিতে column space আর image-এর মধ্যে সম্পর্ক স্থাপিত হয়েছিল, সেই একই রকম যুক্তিতে এটাও দেখা যায় যে--

> কোনো linear transformation-এর kernel থেকে আসে তার matrix-এর nullspace.

সেই একই উদাহরণ নিয়ে ব্যাপারটা বোঝা যাক।

Example 29: আমরা ফের basis নেব $\{1, x, x^2\}$. তাই matrix-টা হবে

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

লক্ষ্য করো কোনো $\vec{v} = ax^2 + bx + c$-এর জন্য $T(\vec{v}) = \vec{0}$ হওয়া মানে $a = b = 0$ হওয়া। তার মানে $\ker(T)$-এর সদস্যরা $0x^2 + 0x + c$ আকারের। এবং সত্যিই $\mathscr{N}(A)$-র সদস্যরাও $(c, 0, 0)$ আকারের। ■

এই ব্যাপারটা যেকোনো basis-এর জন্যই ঘটবে, কারণ যাই basis নাও না কেন, $\vec{0} \in V$ সর্বদাই $(0, ..., 0)$-তেই পরিণত হবে।

Exercise 36: ধরো একটা linear transformation আছে, T. তার domain আর codomain-এর basis-গুলো একভাবে নিলে matrix-টা হয় A, আর অন্য আরেকভাবে নিলে হয় B. তবে কি $\mathscr{N}(A)$ আর $\mathscr{N}(B)$ সমান হতে বাধ্য? ■

27.3 Row space বনাম কী?

Image থেকে আসে column space, আর kernel থেকে আসে nullspace. স্বভাবতঃই প্রশ্ন ওঠে-- row space কীসের থেকে আসে? দুঃখের কথা এর সহজ কোনো উত্তর নেই। পরে আমরা যখন dual space বলে একটা ধারণা আলোচনা করব, তখন এই প্রশ্নের একটা জটিল উত্তর পাওয়া যাবে।

--চার সের আলুর দাম যদি দশ আনা হয়, তবে আধ মণ পটলের দাম কত?
--আহা, আবার পটল এল কোত্থেকে?
--তা তো জানি না। বোধহয় পটলডাঙা থেকে!
--ঝালাপালা (সুকুমার রায়)

27.4 যোগ, বিয়োগ, এবং scalar দিয়ে গুণ

Linear transformation-দের যোগ, বিয়োগ বা কোনো scalar দিয়ে গুণ করলে যে আবার linear transformation-ই পাওয়া যায়, সেটা বোঝা খুবই সহজ। যেমন ধরো, যদি $f, g : V \to W$ দুটো linear transformation হয়, তবে $f+g$ যে ফের V থেকে W-তে একটা linear transformation হবে। সেটা দেখার জন্য যেকোনো $\vec{v}_1, \vec{v}_2 \in V$ নাও আর দুটো scalar নাও a, b. তবে--

$$\begin{aligned}(f+g)(a\vec{v}_1 + b\vec{v}_2) &= f(a\vec{v}_1 + b\vec{v}_2) + g(a\vec{v}_1 + b\vec{v}_2)\\ &= af(\vec{v}_1) + bf(\vec{v}_2) + ag(\vec{v}_1) + bg(\vec{v}_2)\\ &= a(f+g)(\vec{v}_1) + b(f+g)(\vec{v}_2).\end{aligned}$$

Exercise 37: একইভাবে বিয়োগ আর scalar দিয়ে গুণের ক্ষেত্রেও দেখাও। তার মানে যাবতীয় linear transformation-রা মিলে ফের একটা vector space সৃষ্টি করছে! ■

এবার দেখি linear transformation-দের matrix দিয়ে প্রকাশ করলে এইরকম যোগ, বিয়োগ আর scalar দিয়ে গুণের প্রভাব কী হয়। উত্তরটা আন্দাজ করা কঠিন নয়, matrix-গুলোও যোগ, বিয়োগ বা scalar দিয়ে গুণ হয়ে যায়। অবশ্যই domain আর codomain-এর basis-দুটো স্থির রাখতে হবে। ধরো $f, g : V \to W$ দুটো linear transformation আর a, b দুটো scalar. এবার V আর W-এর একটা করে basis নাও। তাদের সাপেক্ষে f আর g-এর matrix ধরো A আর B. তবে সেই একই basis-দের সাপেক্ষে $af + bg$-এর matrix হবে $aA + bB$.
এর মধ্যে কোনোই জটিলতা নেই। জটিলতা যেখানে আছে সেটা হল matrix multiplication-এ। এবার সেটা আলোচনা করব।

27.5 গুণ বনাম composition

Matrix-এর গুণের নিয়মটা যে একটু অদ্ভুত, সেটা সকলেই স্বীকার করবে। এইভাবে গুণ করলে সুবিধা কোথায়, সে বিষয়ে আমরা প্রথম অধ্যায়ে গোটা কয়েক উদাহরণ দিয়েছি বটে, কিন্তু কী করে এই গুণের সংজ্ঞাটা এল, সে বিষয়ে আপাততঃ নীরব থেকেছি। আসলে এটাও এসেছে linear transformation-এর ধারণা থেকেই।
ধরো দুটো linear transformation দিলাম $f : U \to V$ আর $g : V \to W$. তাহলে $g \circ f : U \to W$ একটা linear transformation হতে বাধ্য। সেটা দেখা যায় সহজেই। তার জন্য যা খুশি $\vec{u}_1, \vec{u}_2 in U$ নাও আর যা খুশি দুটো scalar নাও, a, b. তবে--

$$\begin{aligned}(g \circ f)(a\vec{u}_1 + b\vec{u}_2) &= g\big(f(a\vec{u}_1 + b\vec{u}_2)\big)\\ &= g\big(af(\vec{u}_1) + bf(\vec{u}_2)\big)\\ &= ag\big(f(\vec{u}_1)\big) + bg\big(f(\vec{u}_2)\big)\\ &= a(g \circ f)(\vec{u}_1) + b(g \circ f)(\vec{u}_2).\end{aligned}$$

যদি f-এর matrix-টা হয় P আর g-এরটা হয় Q, তবে $g \circ f$-এরটা হবে QP. গুছিয়ে লিখলে পাবে--

> **THEOREM**
>
> Let $f : U \to V$ and $g : V \to W$ be two linear transformations. Let $\mathcal{B}_U, \mathcal{B}_V$ and $\mathcal{B}_W$ be any bases for U, V and W, respectively. If the matrix representing f w.r.t. $\mathcal{B}_U$ and $\mathcal{B}_V$ is P and the matrix representing g w.r.t. $\mathcal{B}_V$ and $\mathcal{B}_W$ is Q, then the matrix representing $g \circ f$ w.r.t. $\mathcal{B}_U$ and $\mathcal{B}_W$ is QP.

প্রমাণের মধ্যে কোনো জটিলতা নেই। শুরু করছি কোনো $\vec{x}$ থেকে, f লাগানোর পর $P\vec{x}$-এ পৌঁছচ্ছি, তার উপর g লাগালে পাচ্ছি $Q(P\vec{x}) = (QP)\vec{x}$. ব্যস্!
Theorem-টার মধ্যে একটা জিনিস লক্ষণীয়। আমরা U থেকে W-তে যাচ্ছি, পথে V-কে ছুঁয়ে। যদি এই V-এর basis-টাকে (মানে $\mathcal{B}_V$-কে) বদলাই, তবে P আর Q বদলে যেতেই পারে, কিন্তু QP একই থাকবে, কারণ $g \circ f$-এর matrix বার করতে $\mathcal{B}_V$ কোনো ভূমিকা পালন করছে না!

27.6 Rank

আমরা দেখেছি যে, একটা linear transformation-এর image থেকেই তার matrix-এর column space-টা আসে। সুতরাং image-এর dimension আর column space-এর dimension যে সমান হবে এটাই স্বাভাবিক, অর্থাৎ linear transformation-এর rank আর তার matrix-এর rank সমান। সেটা নিয়েই নীচের theorem-টা।

> **THEOREM**
>
> Let $T : V \to W$ be a linear transformation. Let A be the matrix of T wrt some bases of V and W. Then rank of $T = r(A)$.

Proof:

> Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be the chosen basis of V, and let $\{\vec{w}_1, ..., \vec{w}_m\}$ be the chosen basis of W.
> Let $A \equiv ((a_{ij}))$.

এইভাবে লিখে নেওয়া মানে এইটা বলা যে A-র entry-রা হল a_{ij}.

> Shall show that rank of $T = r(A)$.
> Let $r = r(A)$.

এবার আমরা $\text{Im}(T)$-এর একটা r সাইজের basis দেব। তবেই প্রমাণ হয়ে যাবে যে, $\dim(\text{Im}(T)) = r$.

> Then, by contraction theorem, there are r columns of A that form a basis of $\mathscr{C}(A)$.
> W.l.g, let these be the first r columns.

এরা প্রথম r-খানা column না হলেও একই যুক্তি খাটত, খালি notation-টা জটিল হয়ে উঠত।

> Shall show that $T(\vec{v}_1), ..., T(\vec{v}_r)$ form a basis of $\text{Im}(T)$.
> This will prove that $r(T) \equiv \dim(\text{Im}(T)) = r$.

Basis দেখানো জন্য দুটো জিনিস দেখাতে হয়--linearly independent আর spanning. প্রথমে linearly independent দেখাব--

> Let $c_1, ..., c_r$ be scalars such that
> $$c_1 T(\vec{v}_1) + \cdots + c_r T(\vec{v}_r) = \vec{0}_W.$$

দেখাতে হবে যে $c_1 = \cdots = c_r = 0$ হতে বাধ্য।

> Now $T(\vec{v}_j) = \sum_i a_{ij} \vec{w}_i$.

এটা আবার কোথা থেকে এল রে, বাবা? আসলে যেহেতু A হল T-এর matrix, তাই $T(\vec{v}_j)$-কে tuple হিসেবে ভাবলে A-র j-th column-টা পাবে, মানে $(a_{1j}, ..., a_{mj})$. সেটা W-এর basis-টার সাপেক্ষে হল

$$a_{1j}\vec{w}_1 + \cdots + a_{mj}\vec{w}_m.$$

ঠিক সেটাই লিখেছিলাম!

> Since $\sum_{j=1}^r c_j T(\vec{v}_j) = \vec{0}_W$, hence
> $$\sum_{j=1}^r \left[c_j \sum_i a_{ij} \vec{w}_i \right] = \vec{0}_W$$
> এবার summation-দুটো আগেপরে করব।
> $$\text{or,} \quad \sum_i \left[\sum_j c_j a_{ij} \right] \vec{w}_i = \vec{0}_W.$$
> $\because \vec{w}_i$'s are linearly independent,
> $$\therefore \forall i \quad \sum_j c_j a_{ij} = 0,$$
> ie,
> $$\sum_{j=1}^r c_j \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$
> $\because$ The first r columns of A are linearly independent,
> $\therefore c_1 = \cdots = c_r = 0$.
> Hence $T(\vec{v}_1), ..., T(\vec{v}_r)$ are linearly independent.

এবার spanning দেখাব, মানে দেখাব যে, যেকোনো $\vec{w} \in \text{Im}(T)$-কেই $T(\vec{v}_1), ..., T(\vec{v}_r)$-এর linear combination হিসেবে লেখা যায়।

> Take any $\vec{w} \in \mathrm{Im}(T)$.
> Then $\therefore \vec{w} = T(\vec{v})$ for some $\vec{v} \in V$.

আমরা $\vec{v}$ আর $\vec{w}$ দুজনকেই tuple হিসেবে লিখব--

> Now $\vec{w} = d_1\vec{w}_1 + \cdots + d_m\vec{w}_m$ for some scalars d_i's,
> and $\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$ for some scalars c_j's.

তার মানে $\vec{w}$-এর tuple হল $(d_1, ..., d_m)$ আর $\vec{v}$-এর tuple হল $(c_1, ..., c_m)$. যেহেতু $\vec{v}$-এর উপর T লাগালে W হয়, তাই $(c_1, ..., c_m)$-কে A দিয়ে গুণ করলে $(d_1, ..., d_m)$ পাবে--

> Then
> $$\begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix} = A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

এইবারে সামান্য কৌশল করব।

> So $(d_1, ..., d_m) \in \mathscr{C}(A)$.
> Let A_1 be the submatrix of A consisting of the first r columns of A which form a basis of $\mathscr{C}(A)$. So $\mathscr{C}(A) = \mathscr{C}(A_1)$, and so $(d_1, ..., d_m) \in \mathscr{C}(A_1)$, ie, it is a linear combination of columns of A_1.

এবারই মোক্ষম ধাপ, যেখানে মাথাটা গুলিয়ে যেতে চাইবে--

> So $\sum d_i\vec{w}_i$ is a linear combination of
> $$\sum_i a_{i1}\vec{w}_i, ..., \sum_i a_{ir}\vec{w}_i,$$

একটা উদাহরণ দিয়ে বোঝা যাক। ধরো $m = r = 2$ আর A-র প্রথম column দুটো $\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$ এবং $\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$. বলে দিয়েছে যে $\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ হল এদের linear combination. ধরো

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 2\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + 3\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}.$$

তবে

$$d_1\vec{w}_1 + d_2\vec{w}_2 = (2a_{11} + 3a_{12})\vec{w}_1 + (2a_{21} + 3a_{22})\vec{w}_2 = 2(a_{11}\vec{w}_1 + a_{21}\vec{w}_2) + 3(a_{12}\vec{w}_1 + a_{22}\vec{w}_2)$$

হবে। ঠিক সেটাই বলেছিলাম।
এবার মনে করে দেখো যে, $\sum_i a_{i1}\vec{w}_i, ..., \sum_i a_{ir}\vec{w}_i$-ই হল $T(\vec{v}_1), ..., T(\vec{v}_r)$.

> which are actually $T(\vec{v}_1), ..., T(\vec{v}_r)$.
> Thus $\vec{w}$ is indeed a linear combination of $T(\vec{v}_1), ..., T(\vec{v}_r)$.
> Hence $\{T(\vec{v}_1), ..., T(\vec{v}_r)\}$ spans $\mathrm{Im}(T)$.

এবার খালি একটা উপসংহার টানার অপেক্ষা--

> Combining, $\{T(\vec{v}_1), ..., T(\vec{v}_r)\}$ is a basis of $\mathrm{Im}(T)$.
> So $\dim(\mathrm{Im}(T)) = r$, as required.

এখানে মোটেই জটিল কিছু হচ্ছে না, খালি notation-এর চচ্চড়ি। ভালো করে বুঝে বুঝে পড়। **[Q.E.D]**

এই theorem-টার সুবাদে এখন কোনো linear transformation-এর rank বার করার জন্য খালি কোনো একজোড়া সুবিধামত basis-এর সাপেক্ষে তার matrix-এর rank বার করলেই হবে। সেটা করা সাধারণতঃ অপেক্ষাকৃত সহজ, কারণ চাইলে RREF ইত্যাদি জোরালো হাতিয়ার লাগানো যেতে পারে।
এবার সেরকম একটা উদাহরণ দেখি।

Example 30: Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(x_1, x_2, x_3) = (2x_1 + x_2 + 3x_3,\ \ 3x_1 - x_2 + x_3,\ \ -4x_1 + 3x_2 + x_3).$$

Find rank of T.
SOLUTION:

> Let A be the matrix representing this linear transformation wrt the standard ordered basis. Then for all $(x_1, x_2, x_3) \in \mathbb{R}^3$
>
> $$T(x_1, x_2, x_3) = A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 3 & -1 & 1 \\ -4 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$
>
> Thus,
>
> $$A = \begin{bmatrix} 2 & 1 & 3 \\ 3 & -1 & 1 \\ -4 & 3 & 1 \end{bmatrix}.$$

এর rank বার করার নানা কায়দা আছে। RREF বার করতে পারো, বা সরাসরি গায়ের জোরেও করতে পারো, যেমন--

> Here, by inspection, the first two rows are not multiples of each other.
> So $r(A) \geq 2$.

এবার দেখতে হবে তৃতীয় row-টাকে এদের linear combination হিসেবে লেখা যায় কিনা--

> Let, if possible, $(-4, 3, 1) = a(2, 1, 3) + b(3, -1, 1)$.
> Then $2a + 3b = -4$, $a - b = 3$ and $3a + b = 1$. From the last two, we get $a = 1, b = -2$, which satisfy the first equation.
> So $r(A) = 2$.

Hence the rank of T is 2.

অবশ্যই একটু বড় matrix হলে এরকম পথে এগোনো দুষ্কর। তখন হাতের পাঁচ $RREF$-ই ভরসা। ■

এবার একটা ধাঁধানো প্রশ্ন।

Exercise 38: ধরো V হল যাবতীয় $a_0 + a_1x + \cdots + a_{100}x^{100}$ চেহারার polynomial-দের vector space, যেখানে a_i-রা সবাই real. তবে differentiation হল V থেকে V-তে একটা linear transformation. এর rank কত?
HINT: কোনো RREF বা ভারী কিছু যন্ত্রপাতি লাগবে না, খালি একটু চিন্তা করতে হবে, ব্যস।! ■

27.7 Nullity

অনুমান করতে পারছ যে, rank-এর মত nullity-র বেলাতেও linear transformation-এর kernel-এর dimension হবে (যেকোনো একজোড়া basis-এর সাপেক্ষে) তার matrix-এর nullity. সেটা গুছিয়ে লিখে প্রমাণ করার দায়িত্ব তোমাকেই দিলাম। ঠিক যেভাবে আগের theorem-টা প্রমাণ হল, সেইভাবে এগোলেই হবে। তবে যারা প্রথমবার পড়ছ, তারা এখনই প্রমাণের জটিলতায় যেও না। আগে এই অধ্যায়ের বাকিটা শেষ করে নাও।

DAY 28
Rank + nullity = domain dimension

এবার rank আর nullity-র মধ্যে একটা অদ্ভুত সম্পর্ক শিখব। ওদের যোগ করলে সর্বদাই domain-এর dimension-টা পাওয়া যায়! ব্যাপারটা বিশেষভাবে অদ্ভুত এই কারণে যে, rank হল image-এর dimension, আর nullity হল kernel-এর। এই যে দুটো subspace, এরা রয়েছে সম্পূর্ণ আলাদা দুটো vector space-এর মধ্যে--image আছে codomain-এর মধ্যে, আর kernel আছে domain-এর মধ্যে। সুতরাং এই আলাদা দুই জগতের subspace দুটোর union, sum বা intersection কিছুই নেওয়া সম্ভব নয়। ওদের dimension দুটোকে খামোখা যোগ করার পিছনে কোনো সহজ ব্যাখ্যাও পাওয়া যায় না। কিন্তু তাও কি আশ্চর্য, যোগফলটা domain-এর dimension হয়ে যায়! এই ব্যাপারটা অংকের ভাষায় গুছিয়ে লিখে প্রমাণ করার আগে কয়েকটা উদাহরণ দেখে নিই।

28.1 কিছু উদাহরণ

Example 31: Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ which maps the basis vectors $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ of $\mathbb{R}^3$ to the vectors $(2, 0, 0)$, $(0, 2, 0)$, $(0, 0, 2)$ respectively. Find $\ker(T)$ and $\mathrm{Im}(T)$. Verify that $\dim(\ker(T)) + \dim(\mathrm{Im}(T)) = 3$.[2+1+1+1] **(2019)**
SOLUTION:এই অংকের প্রথম তিনটে অংশ আমরা আগেই করেছি Example 24-এ (সেটা পাবে 216 নম্বর পৃষ্ঠার ধারেকাছে)। সেখানে আমরা দেখেছিলাম যে, $\ker(T) = \{(0, 0, 0)\}$ এবং $\mathrm{Im}(T) = \mathbb{R}^3$. তাই $\dim(\ker(T)) = 0$ আর $\dim(\mathrm{Im}(T)) = 3$. সত্যিই ওদের যোগ করলে 3 হচ্ছে, এবং এখানে domain হল $\mathbb{R}^3$, যার dimension-ও 3. ■

Exercise 39: Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ that maps the basis vectors $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ of $\mathbb{R}^3$ to the vectors $(2, 1, 1)$, $(1, 2, 1)$, $(1, 1, 2)$ respectively. Find $\ker(T)$ and $\mathrm{Im}(T)$. Verify that $\dim(\ker(T)) + \dim(\mathrm{Im}(T)) = 3$.[2+1+1+1] **(2020.CC6.3a)**
HINT: এই অংকের গোড়ার দিকটা Exercise 33-এ করতে দিয়েছিলাম (217 নম্বর পাতার কাছাকাছি)। সেটা যদি করে থাকো, তবে এ অংকটা এক মুহূর্তে হয়ে যাবে। ■

কেন rank আর nullity যোগ করলে domain-এর dimension হয়, সেটা এবার আমরা অংকের ভাষায় গুছিয়ে লিখে প্রমাণ করব।

28.2 অংকের ভাষায়

> **THEOREM**
>
> Let V, W be two vector spaces both defined on the same field, F, where V is finite dimensional. Let $T : V \to W$ be any linear transformation. Then
>
> $$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

লক্ষ্য কর, এখানে আমরা W-কে finite dimensional ধরিনি, কারণ তার কোনো দরকার নেই।

Proof: ধরো $\text{nullity}(T) = n$ আর $\text{rank}(T) = r$. আমাদের দেখাতে হবে $\dim(V) = r + n$ হবে। সেটা দেখানোর জন্য একদম সোজা কায়দা হল এটা দেখানো যে, V-এর একটা $r+n$ সাইজের basis আছে। আমরা এরকম একটা basis একেবারে হাতেকলমে তৈরী করে দেখিয়ে দেব। তার মানে আমাদের হাতে এখন তিনটে কাজ--

- এক, V-র একটা $r+n$ সাইজের subset নেওয়া,
- দুই, সেই subset-টাকে linearly independent দেখানো,
- আর তিন, এটা দেখানো যে subset-টা V-কে generate করে।

একে একে কাজগুলো সেরে ফেলা যাক--

> Step 1: Let $\{\vec{w}_1, ..., \vec{w}_r\}$ be any basis of $\text{Im}(T)$, where $r = \text{rank}(T) = \dim(\text{Im}(T))$.
>
> Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be any basis of $\ker(T)$, where $r = \text{nullity}(T) = \dim(\ker(T))$.

বাঃ এই তো r-খানা $\vec{w}_i$ আর n-খানা $\vec{v}_i$ পাওয়া গেছে, এদেরকে একসাথে মিলিয়ে তো বেশ একটা $r+n$ সাইজের set হতে পারে! সেটাই তো দরকার ছিল! কিন্তু সমস্যা হল, $\vec{w}_i$-গুলো আছে W-এর মধ্যে। আমাদের set-টাকে তো V-এর subset হতে হবে। সুতরাং কোনোভাবে $\vec{w}_i$-গুলো ব্যবহার করে V-এর কিছু vector পেতে হবে।

> Now, $\because \vec{w}_1, ..., \vec{w}_r \in \text{Im}(T)$,
>
> $\therefore \exists \vec{u}_1, ..., \vec{u}_r \in V \quad \forall i \quad T(\vec{u}_i) = \vec{w}_i$.

এইবার তবে $\vec{v}_i$ আর $\vec{u}_i$-গুলোকে মিলিয়ে একটা set বানানো যাক। সেটা অবশ্যই V-এর subset হবে। প্রশ্ন হল সাইজটা $r+n$ হবে তো? একটা সমস্যা হতে পারে-- যদি $\vec{u}_i$ আর $\vec{v}_i$-রা সবাই আলাদা আলাদা না হয়? কিন্তু এখানে সে সমস্যা হবে না, কারণ--

> Since $\{\vec{w}_1, ..., \vec{w}_r\}$ is a basis, so they are all distinct. So $\vec{u}_1, ..., \vec{u}_n$ are also all distinct.
>
> Similarly, being elements of a basis, $\vec{v}_1, ..., \vec{v}_n$ are all distinct.

Also the $\vec{u}_i$'s are distinct from the $\vec{v}_j$'s.

এটা কেন হবে বোঝা যাক। $T(\vec{u}_i) = \vec{w}_i$, আর $\vec{w}_i \neq \vec{0}_W$, কারণ $\vec{w}_i$-রা মিলে একটা linearly independent set হয়। কিন্তু $T(\vec{v}_j) = \vec{0}_W$. সেই কারণে কোনো $\vec{u}_i$ কোনো $\vec{v}_j$-র সমান হতে পারে না।

⟦ Because:

$\forall i \quad T(\vec{u}_i) = \vec{w}_i \neq \vec{0}_W$, but $\forall j \quad T(\vec{v}_j) = \vec{0}_W$.

⟧

Thus $B = \{\vec{u}_1, ..., \vec{u}_r, \vec{v}_1, ..., \vec{v}_n\} \subseteq V$ has size $r + n$.

তাহলে $r + n$ সাইজের একটা subset পাওয়া গেল V-এর মধ্যে। এবার দেখাতে হবে যে, এটা আসলে V-এর একটা basis. প্রথমে linearly independent দেখাই--

Step 2: Shall show that B is linearly independent.

Let, for some $\alpha_1, ..., \alpha_r, \beta_1, ..., \beta_n \in F$,

$$\alpha_1\vec{u}_1 + \cdots + \alpha_r\vec{u}_r + \beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n = \vec{0}_V. \qquad (*)$$

আমরা দেখাব যে, এর ফলে α_i আর β_j-রা সবাই শূন্য হতে বাধ্য।

Applying T to both sides,

$$T(\alpha_1\vec{u}_1 + \cdots + \alpha_r\vec{u}_r + \beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n) = T(\vec{0}_V).$$

Hence, by linearity of T,

$$\alpha_1 T(\vec{u}_1) + \cdots + \alpha_r T(\vec{u}_r) + \beta_1 T(\vec{v}_1) + \cdots + \beta_n T(\vec{v}_n) = \vec{0}_W,$$

or

$$\alpha_1\vec{w}_1 + \cdots + \alpha_r\vec{w}_r = \vec{0}_W,$$

since $T(\vec{u}_i) = \vec{w}_i$ and $T(\vec{v}_j) = \vec{0}_W$.

But $\{\vec{w}_1, ..., \vec{w}_r\}$ is a linearly independent set. So

$$\alpha_1 = \cdots = \alpha_r = 0.$$

যাক α_i-গুলোকে সংহার করা গেছে! এবার β_i-দেরও শূন্য দেখাব--

Hence, from (*),

$$\beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n = \vec{0}_V.$$

But $\{\vec{v}_1, ..., \vec{v}_n\}$ is a linearly independent set. So

$$\beta_1 = \cdots = \beta_n = 0.$$

This completes the proof of linear independence.

আর খালি একটাই কাজ পড়ে আছে, generating set দেখানো।

Step 3: Shall show that B spans V, ie,

$\forall \vec{v} \in V \quad \exists \alpha_1, ..., \alpha_r, \beta_1, ..., \beta_n \in F \quad \vec{v} = \sum_{i=1}^{r} \alpha_i \vec{u}_i + \sum_{j=1}^{n} \beta_j \vec{v}_j$.

সাবধান, এই α, β-দের সঙ্গে আগের ধাপের α, β-দের কোনো সম্পর্ক নেই কিন্তু!

$\forall \vec{v}$ Take any $\vec{v} \in V$.

এবার α_i, β_j-গুলো চাই--

Let $\vec{w} = T(\vec{v})$.
$\because \vec{w} \in W$ and $\{\vec{w}_1, ..., \vec{w}_r\}$ is a basis of W,
$\therefore \exists \alpha_1, ..., \alpha_r \in F \quad \vec{w} = \sum_i \alpha_i \vec{w}_i$.
$\exists \alpha_i$ Choose these $\alpha_1, ..., \alpha_r$.

α_i-দের পাওয়া গেল। এবার β_j-দের জোগাড় করতে হবে--

Then $T\left(\sum_i \alpha_i \vec{u}_i\right) = \sum_i \alpha_i \vec{w}_i = \vec{w}$.
Also we had $T(\vec{v}) = \vec{w}$.
$\therefore T\left(\vec{v} - \sum_i a_i \vec{u}_i\right) = \vec{w} - \vec{w} = \vec{0}_W$.
Thus, $\vec{v} - \sum_i a_i \vec{u}_i \in \ker(T)$.
Since $\{\vec{v}_1, ..., \vec{v}_n\}$ is a basis of $\ker(T)$,
So $\exists \beta_1, ..., \beta_n \in F \quad \vec{v} - \sum_i a_i \vec{u}_i = \sum_j \beta_j \vec{v}_j$.
$\exists \beta_j$ Choose these β_j's.

ব্যস্, β_j-রাও উপস্থিত। সামান্য কাজই এখন বাকি।

Then

$$\vec{v} = \sum_i a_i \vec{u}_i + \sum_j \beta_j \vec{v}_j.$$

এবার একটা উপসংহার--

So $\{\vec{u}_1, ..., \vec{u}_r, \vec{v}_1, ..., \vec{v}_n\}$ is a basis of size $r + n$ for V.
Hence $\dim(V) = r + n$, as required.

[Q.E.D]

এই theorem-টায় W-এর finite dimensional হওয়ার কোনো দরকার ছিল না। কিন্তু যদি তা হত, তাহলে আমরা T-কে একটা matrix দিয়ে লিখতে পারতাম (V আর W-এর কোনো basis-এর সাপেক্ষে)। আন্দাজ করতে পারছ যে, theorem-টাকে আমরা matrix-টা দিয়েও দিয়ে প্রকাশ করতে পারি। যদি matrix-টাকে A নাম দিই, তবে $\text{rank}(T) = r(A)$ আর $\text{nullity}(T) = \text{nullity}(A)$. কিন্তু $\dim(V)$ এই matrix-এর ভাষায় কী? উত্তর হল, $\dim(V)$ হচ্ছে A-র column-সংখ্যা। তাই theorem-টা বলছে--

$r(A) + \text{nullity}(A) = A$-র column-সংখ্যা।

যেহেতু যেকোনো matrix-ই কোনো না কোনো linear transformation-কে প্রকাশ করে, তাই এ জিনিসটা যেকোনো matrix-এর বেলাতেই খাটবে। ফলে পাব নীচের theorem-টা--

THEOREM

For any $A_{\sqcup imesn}$ we have

$$r(A) + \text{nullity}(A) = n.$$

28.3 হাতে কলমে

আমরা একটু আগের কয়েকটা উদাহরণের rank আর nullity আলাদা করে বার করে যোগ করে দেখেছিলাম যে domain-এর dimension-এর সমান হচ্ছে। এবার যখন Theorem-টা থেকে গ্যারান্টি পেয়ে গেছি যে, এমনটা সবসময়েই হবে, আমরা rank আর nullity বার করার সময়ে বেশ কিছুটা খাটনি বাঁচাতে পারি। Rank আর nullity-র মধ্যে যেটা বেশী সহজ লাগে, সেটা বার করে, domain -এর dimension থেকে বিয়োগ করে দিলেই অন্যটা বেরিয়ে যাবে। যেমন ধরো যদি কোনো $T : \mathbb{R}^4 \to \mathbb{R}^4$-এর জন্য $\text{nullity}(T) = 0$ পেয়ে যাই, অমনি তৎক্ষণাৎ $\text{rank}(T) = 4 - 0 = 4$ বলে দিতে পারব। শুধু তাই নয়, যেহেতু codomain-টা $\mathbb{R}^4$, তাই এটাও বলে দিতে পারব যে $\text{Im}(T) = \mathbb{R}^4$ হতে বাধ্য!
এই কায়দায় কয়েকটা অংক করি।

Example 32: If $T : \mathbb{R}^4 \to \mathbb{R}^3$ is defined as $T(x, y, z, t) = (x - y - z - t, x + 2z - t, x + y + 3z - 3t)$,

then find the rank and nullity of T.[3+2] **(2021)**
SOLUTION: এই অংকটা আমরা আগেই দেখেছি। 212 নম্বর পাতার ধারেকাছে Example 20-তে আমরা বার করেছিলাম $\text{rank}(T) = 3$. যেহেতু এখানে domain হল $\mathbb{R}^4$, যার dimension হচ্ছে 4, তাই $\text{nullity}(T) = 4 - 3 = 1$. ■

Example 33: Determine the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^4$ that maps the vectors $(1, 2, 3), (1, 3, 2), (2, 3, 1)$ of $\mathbb{R}^3$ to the vectors $(0, 1, 1, 1), (1, 0, 1, 1)$ and $(1, 1, 0, 1)$, respectively. Find $\ker(T)$ and $\text{rank}(T)$.
SOLUTION:

Given any $(x, y, z) \in \mathbb{R}^3$, we try to express it as

$$(x, y, z) = a(1, 2, 3) + b(1, 3, 2) + c(2, 3, 1),$$

or

$$\begin{aligned} a+b+2c &= x \\ 2a+3b+3c &= y \\ 3a+2b+c &= z. \end{aligned}$$

Solving

$$\begin{aligned} a &= \tfrac{1}{2}(x-y+z) \\ b &= \tfrac{1}{6}(-7x+5y-z) \\ c &= \tfrac{1}{6}(5x-y-z). \end{aligned}$$

Thus, a, b, c are uniquely determined by x, y, z.

$$\begin{aligned} \therefore T(x,y,z) &= T(a(1,2,3)+b(1,3,2)+c(2,3,1)) \\ &= aT(1,2,3)+bT(1,3,2)+cT(2,3,1) \\ &= a(0,1,1,1)+b(1,0,1,1)+c(1,1,0,1) \\ &= (b+c, a+c, a+b, a+b+c) \\ &= \tfrac{1}{3}\left(-x+2y-z,\ 4x-2y+z,\ -2x+y+z,\ \tfrac{1}{2}(x+y+z)\right), \end{aligned}$$

which is the required linear transformation.

এবার kernel বার করি--

Setting $T(x,y,z) = (0,0,0,0)$, and solving, we get $x = y = z = 0$.
So $\ker(T) = \{(0,0,0)\}$.

সুতরাং $\ker(T)$ দেখা যাচ্ছে $\mathbb{R}^3$-এর trivial subspace. অতএব--

$\therefore \text{nullity}(T) = 0.$

এইবার আমাদের theorem-টা লাগাব। মনে রেখো এখানে আমাদের domain-টা হল $\mathbb{R}^3$.

$\therefore \text{rank}(T) = \dim(\mathbb{R}^3) - \text{nullity}(T) = 3 - 0 = 3.$

■

এবার তোমার করার জন্য কিছু অংক দিই।

Exercise 40: Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be defined by

$$T(x_1, x_2, x_3, x_4) = (3x_1 - 2x_2, -x_3 - x_4,\ x_1 + x_2 - 2x_3 - 3x_4),$$

where $(x_1 x_2, x_3, x_4) \in \mathbb{R}^4$. Find $\text{rank}(T)$, $\text{nullity}(T)$ and a basis of $\ker(T)$.
HINT:

যদিও এখানে প্রথমে rank, তারপর nullity এবং সবশেষে $\ker(T)$-এর basis বার করতে বলেছে, কিন্তু আসলে উল্টো দিক দিয়ে এগোলেই বেশী সুবিধা হবে। আমরা তাই শুরু করব $\ker(T)$-এর একটা basis বার করে, basis-এ কতগুলো vector আছে সেটা গুণে নিলেই $\text{nullity}(T)$ বেরিয়ে যাবে, তারপর বিয়োগ করে rank বার করতে আর কতক্ষণ? ■

Exercise 41: A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by

$$T(x_1, x_2.x_3) = (x_1 + x_2 + x_3,\ \ 2x_1 + x_2 + 2x_3,\ \ x_1 + 2x_2 + x_3).$$

Find $\ker(T)$ and dimensions of $\ker(T)$ and $\text{Im}(T)$. ■

DAY 29
Injective, surjective, bijective

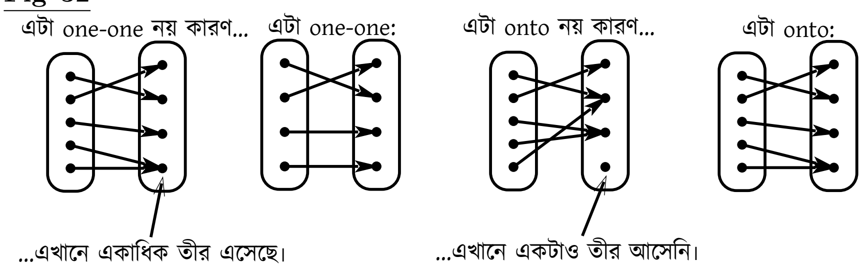
29.1 One-one, onto, bijection

Linear transformation-রা এক ধরণের function. তাই যেকোনো function-এর মত এরাও one-one বা onto হতে পারে। বা দুটোই একই সঙ্গে হতে পারে, বা কোনোটাই নাও হতে পারে। One-one বা onto (যাদের পোশাকী নাম যথাক্রমে **injective** এবং **surjective**) কাদের বলে যদি ভুলে গিয়ে থাকো, তবে Fig 32 দেখলেই মনে পড়ে যাওয়া উচিত। একটা function যদি একই সঙ্গে injective এবং surjective হয়, তবে তাকে বলে **bijective** বা **invertible**.
Linear transformation-দের বেলায় এদের আরেক প্রস্থ করে বাহারি নাম আছে, যদিও সেগুলো তত ব্যবহার হয় না। Linear transformation-দের অনেক সময়ে বলে **homomorphism of vector spaces.** যদি সেটা injective হয়, তবে তাকে বলে **monomorphism of vector spaces.** আর surjective হলে তার নাম **epimorphism of vector spaces.** আর দুটোই এক সঙ্গে হলে তাকে বলে **isomorphism of vector spaces.**
এই নামগুলো linear algebra শেখার জন্য কিছুমাত্র প্রয়োজনীয় নয়। কিন্তু যারা খানিকটা group theory শিখেছ (যেমন আমাদের বাংলায় বোঝানো ইংরাজি বই সিরিজের Modern Algebra থেকে), তারা লক্ষ্য করবে যে, group-এর homomorphism-এর সঙ্গে vector space-এর homomorphism-এর অনেক মিল আছে। যদি group theory না জানা থাকে, তবে দুশ্চিন্তায় পড়ে যেও না। এই বইটা বোঝার জন্য তার বিশেষ কোনো প্রয়োজন হবে না।

Example 34: তিনটে linear transformation দিচ্ছি--

1. $f : \mathbb{R}^2 \to \mathbb{R}$, যেখানে $f(x, y) = x$,
2. $g : \mathbb{R}^2 \to \mathbb{R}^2$, যেখানে $g(x, y) = (x, 0)$,

Fig 32



3. $h : \mathbb{R}^2 \to \mathbb{R}^2$, যেখানে $h(x, y) = (x + y, x)$.

এদের মধ্যে কারা one-one, কারা onto এবং কারা কোনোটাই নয়?

SOLUTION:

1. এখানে f-টা one-one নয়, কারণ y-এর কোনো প্রভাব $f(x, y)$-এর উপর পড়ছে না, তাই $f(0, 1)$ আর $f(0, 100)$ দুজনেই একই--
$$f(0, 1) = f(0, 100) = 0.$$
কিন্তু f অবশ্যই onto. তুমি যাই $x \in \mathbb{R}$ নাও না কেন, আমি তারপরে যা খুশি y জুড়ে দেব, তাহলেই $f(x, y) = x$ হবে, যেমন $\forall x \in \mathbb{R} \quad f(x, 100) = x$.

2. এক্ষুণি f-এর বেলায় যা বললাম, সেই একই কারণে g-ও one-one নয়। কিন্তু g-টা onto-ও নয়। যদি তোমাকে $(1, 1) \in \mathbb{R}^2$ দিই, তবে এমন কোনো $(x, y) \in \mathbb{R}^2$ পাবে না, যাতে $g(x, y) = (1, 1)$ হয়, কারণ $g(x, y)$-এর দ্বিতীয় স্থানে 0 থাকতেই হবে।

3. এবার h-এর দিকে তাকাই। তুমি যা খুশি $(a, b) \in \mathbb{R}^2$ নাও। তাহলে কি এমন $(x, y) \in \mathbb{R}^2$ পাবেই, যাতে $h(x, y) = (a, b)$ হয়? মানে $x + y = a$ আর $x - y = b$ হয়? সমাধান করলেই দেখবে যেকোনো (a, b)-র জন্যই সমাধান আছে এবং খালি একটা করেই সমাধান আছে। তাই h-টা একাধারে one-one এবং onto.

■

Exercise 42: এমন একটা linear transformation দাও $f : \mathbb{R} \to \mathbb{R}^2$ যেটা one-one. ■

Exercise 43: এমন একটা linear transformation দাও $f : \mathbb{R} \to \mathbb{R}^2$ যেটা onto. ■

Exercise 44: এমন একটা linear transformation দাও $f : \mathbb{R}^2 \to \mathbb{R}$ যেটা one-one. ■

Exercise 45: এমন একটা linear transformation দাও $f : \mathbb{R}^2 \to \mathbb{R}$ যেটা onto নয়। ■

Example 35: Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear mapping defined by $T(x, y, z) = (x - y, x + 2y, y + 3z)$, $(x, y, z) \in \mathbb{R}^3$. Show that T is invertible. Find T^{-1}.[3+2] **(2019.CC6.3f)**

SOLUTION: Invertible মানে এইভাবে চিন্তা করো--যদি তোমাকে যা খুশি $T(x, y, z)$ বলে দিই, তবে অমনি তুমি নিশ্চিতভাবে (x, y, z) বার করে ফেলতে পারবে। যদি কোথাও আটকে যাও (মানে সেরকম (x, y, z) পাওয়া যাচ্ছে না), তবে onto নিয়ে সমস্যা হবে, আর যদি নিশ্চিত না হতে পারো (মানে একাধিক (x, y, z) বেরোয়), তবে one-one হবে না।

> Let $T(x, y, z) = (p, q, r)$ be given where (p, q, r) is any vector in $\mathbb{R}^3$.
> Then
> $$\begin{aligned} x - y &= p, \\ x + 2y &= q, \\ y + 3z &= r. \end{aligned}$$

এবার এখানে থেকে x, y, z-কে প্রকাশ করার চেষ্টা করব p, q, r দিয়ে। সেটা রাফে করে নিলে পাবে--

> Solving we get $x = \frac{2p+q}{3}$, $y = \frac{q-p}{3}$ and $z = \frac{p-q+r}{3}$.

ব্যস্, যেকোনো $(p, q, r) \in \mathbb{R}^3$-এর জন্যই যখন নিশ্চিতভাবে (x, y, z) বার করে দেওয়া গেল, তখন T-এর invertible হওয়া নিয়ে আর সন্দেহের অবকাশ নেই।

> Since we get unique solution for all $(p, q, r) \in \mathbb{R}^3$, we see that T is invertible.

আর T^{-1}-ও বেরিয়ে গেছে--

> Also $T^{-1}(p, q, r) = (x, y, z) = \left(\frac{2p+q}{3}, \frac{q-p}{3}, \frac{p-q+r}{3}\right)$.

■

Exercise 46: Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T(x, y, z) = (3x + z, \quad -2x + y, \quad -x + 2y + 4z).$$

Examine whether T is invertible. ■

একটা function যদি invertible হয়, তবে তার inverse থাকে। এই inverse-টাও একটা function. এইটুকু যেকোনো invertible function-এর জন্যই খাটে। কিন্তু function-টা যদি একটা linear transformation হয়, তবে একটা বাড়তি বোনাস আছে--এক্ষেত্রে inverse-টাও ফের একটা linear transformation-ই হয়। সেটাই বলেছে নীচের theorem-এ।

THEOREM

Let $T : V \to W$ be an invertible linear transformation. Then $T^{-1} : W \to V$ is again a linear transformation.

Proof:

> Let the field be F.
> Now
>
> $$\forall \vec{v} \in V \quad T^{-1}(T(\vec{v})) = \vec{v},$$
> $$\forall \vec{w} \in W \quad T(T^{-1}(\vec{w})) = \vec{w}.$$
>
> Shall show that
>
> ◎ $\forall a, b \in F \quad \forall \vec{w}_1, \vec{w}_2 \in W \quad T^{-1}(aw_1 + bw_2) = aT^{-1}(\vec{w}_1) + bT^{-1}(\vec{w}_2)$.
>
> [$\forall a, b$] Take any $a, b \in F$.
>
> [$\forall \vec{w}_1, \vec{w}_2$] Take any $\vec{w}_1, \vec{w}_2 \in W$.

∎

এখানে একটা ছোট্টো কৌশল করব। আমাদের দেখাতে হবে $T^{-1}(aw_1 + bw_2)$ আর $aT^{-1}(\vec{w}_1) + bT^{-1}(\vec{w}_2)$ হল সমান। আমরা দুজনের উপরই T লাগাব। তাতে যদি দুই ক্ষেত্রেই একই উত্তর আসে, অমনি বলতে পারব যে ওরা দুজন সমান, কারণ T-টা injective.

Then $a\vec{w}_1 + b\vec{w}_2 \in W$.

$$\begin{aligned} \therefore T(T^{-1}(a\vec{w}_1 + b\vec{w}_2)) &= a\vec{w}_1 + b\vec{w}_2 \\ &= aT(T^{-1}(\vec{w}_1)) + bT(T^{-1}(\vec{w}_2)) \\ &= T(aT^{-1}(\vec{w}_1) + bT^{-1}(\vec{w}_2)). \end{aligned}$$

But T is invertible, and hence injective. So $T^{-1}(a\vec{w}_1 + b\vec{w}_2) = aT^{-1}(\vec{w}_1) + bT^{-1}(\vec{w}_2)$, as required.

[Q.E.D]

29.2 Image এবং rank-এর সাথে surjective-এর সম্পর্ক

একটা linear transformation $T : V \to W$ দিয়ে যদি জিজ্ঞাসা করি সেটা surjective বা injective কিনা, তবে $\mathrm{Im}(T)$ বা $\ker(T)$-র দিকে এক ঝলক তাকিয়েই সেটা বলে দেওয়া যায়। $\mathrm{Im}(T)$ দেখে surjective কিনা বলা অতি সহজ, তার জন্য linear transformation হওয়ারও কোনো দরকার নেই।

যদি $\mathrm{Im}(T) = W$ হয় তবেই T-টা surjective হবে।

এটা বস্তুতঃ surjective-এর সংজ্ঞা ছাড়া আর কিছুই নয়। তাই আলাদা করে প্রমাণ করার কিছু নেই। চাইলে শর্তটাকে আমরা dimension দিয়ে লিখতে পারতাম। সব মিলিয়ে পেতাম এই theorem-টা--

THEOREM

Let $T : V \to W$ be a linear transformation. Then the following are equivalent:

1. T is surjective.
2. $\mathrm{Im}(T) = W$
3. $\mathrm{rank}(T) = \dim(W)$.

এখানে "equivalent" বলতে "সমার্থক" বোঝানো হচ্ছে, মানে এই তিনটে শর্তের মধ্যে একটাও যদি পালিত হয়, তবে অন্য দুটোও পালিত হতে বাধ্য। দ্বিতীয় আর তৃতীয় শর্ত দুটো কেন equivalent বুঝলে তো? যেহেতু $\mathrm{Im}(T)$ হল W-এর একটা subspace, তাই $\dim(\mathrm{Im}(T)) = \dim(W)$ হওয়া মানে $\mathrm{Im}(T) = W$ হওয়া, আর $\dim(\mathrm{Im}(T))$-কেই তো $\mathrm{rank}(T)$ বলে!

29.3 Kernel এবং nullity-র সাথে injective-এর সম্পর্ক

Surjective হওয়ার সঙ্গে Im(T)-এর সম্পর্কটা বোঝা তেমন জটিল কিছু ছিল না। কিন্তু injective হওয়া বা না হওয়ার সঙ্গে $\ker(T)$-এর কী সম্পর্ক সেটা ততটা স্পষ্ট নয়। একটা উদাহরণ নিয়ে দেখি। Fig 33 এবং Fig 34-এ দুটো function

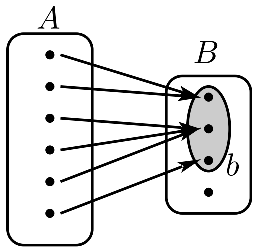


Fig 33

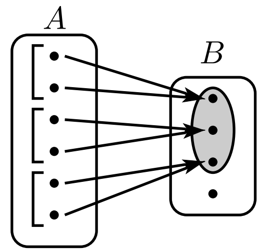


Fig 34

দেখানো আছে। এরা কেউই injective নয়, কারণ দুই ক্ষেত্রেই অন্ততঃ দুটো তীর আছে যারা একই জায়গায় গেছে। কিন্তু Fig 34-এ একটা সৌন্দর্য আছে। সেটা হল তীরগুলো সমান সাইজের কয়েকটা গুচ্ছ সৃষ্টি করেছে, এবং একই গুচ্ছের সব তীরই একই জায়গায় গেছে। তাই image-এর প্রতিটা বিন্দুতেই সমানসংখ্যক তীর এসে মিশেছে। এরকম ক্ষেত্রে function-টা injective কিনা, সেটা image-এর যেকোনো একটা বিন্দুতে কটা তীর এসে মিশেছে, সেটা দেখেই বলে দেওয়া যায়--যদি খালি একটাই আসে, তবে injective, নইলে নয়। অনেকটা হাঁড়ির একটা ভাত টিপেই বলে দিতে পারার মত যে, বাকি ভাতগুলোও সিদ্ধ হয়েছে কিনা। Fig 33-এর বেলায় সেটা সম্ভব নয়, যেমন b-তে একটাই তীর এসেছে, কিন্তু খালি সেটুকু দেখেই injective ভেবে নিলে ভুল করা হবে। Linear transformation-রা হল Fig 34-এর মত সুন্দর জিনিস। সেটা একটা সহজ উদাহরণ দেখলেই বোঝা যাবে।

Example 36: ধরো $T:\mathbb{R}^2\to\mathbb{R}$ হল $T(x_1,x_2)=x_1-x_2$. তোমার কাজ হল $\mathbb{R}^2$-র মধ্যে সেইসব বিন্দুগুলোকে ছবি এঁকে দেখানো, যেখানে $T(x_1,x_2)=0$ হয়। তারপর একই ছবিতে সেই সব বিন্দুগুলোকে দেখাও যেখানে $T(x_1,x_2)=1$ এবং $T(x_1,x_2)=2$ হয়।

SOLUTION: $T(x_1,x_2)=0$ মানে $x_1=x_2$. তার মানে ছবিতে হবে Fig 35-এর মত। যদি $T(x_1,x_2)=1$-এর ছবি আঁকি, তবে সেটা হবে $x_1-x_2=1$, মানে $x_2=x_1-1$-এর গ্রাফ। একইভাবে $T(x_1,x_2)=2$ হবে $x_2=x_1-2$-এর গ্রাফ। ■

এখানে প্রত্যেকটা লাইন যেন একেকটা গুচ্ছ, একই লাইনের উপর বিন্দুরা সবাই একই value-তে যায়। লক্ষ্য করো, লাইনগুলো সব একইরকম দেখতে, খালি খানিকটা করে সরানো। এর মধ্যে যে লাইনটা $(0,0)$ দিয়ে গেছে সেটাই হল $\ker(T)$.
আরেকটা উদাহরণ দেখি।

Example 37: এবার $T:\mathbb{R}^3\to\mathbb{R}$ নেব $T(x_1,x_2,x_3)=x_3$. তাহলে $\mathbb{R}^3$-র কোন্ কোন্ বিন্দু $0,1$ আর 2-তে যাবে?

SOLUTION: ছবিতে দেখাবে Fig 36-র মত। ■

এখানেও দ্যাখো সবগুলো set-ই কেমন একই রকম দেখতে, সবাই একেকটা plane, খালি খানিকটা করে সরানো। যে plane-টা $(0,0,0)$ দিয়ে যায়, সেটাই এখানে $\ker(T)$.

Fig 35

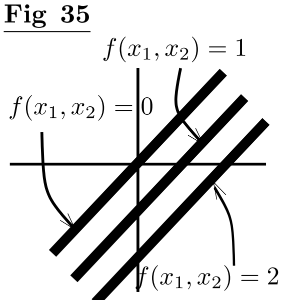
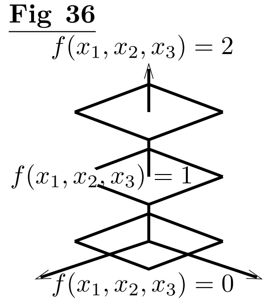


Fig 36



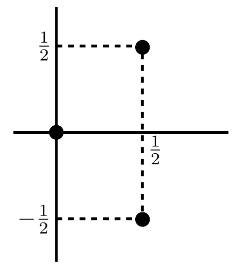


Fig 37

Example 38: যদি $T:\mathbb{R}^2\to\mathbb{R}^2$ নিই $T(x_1,x_2)=(x_1+x_2,x_1-x_2)$, তবে সেই সব বিন্দুর set বার কর, যাতে $T(x_1,x_2)=(0,0)$ হয়। কোন্ কোন্ বিন্দুতে $T(x_1,x_2)=(0,1)$ আর $T(x_1,x_2)=(1,0)$ হবে?
SOLUTION: এই তিনটে set-ই হল একেকটা বিন্দু নিয়ে তৈরী। Fig 37 দ্যাখো। এখানে $\ker(T)=\{(0,0)\}$. ■

যদি কোনো linear transformation $T:V\to W$-এর জন্য দ্যাখো যে $\ker(T)$ হল একটা singleton set মানে খালি একটাই তীর পৌঁছেছে $\vec{0}_W$-তে, অমনি জানবে যে $\text{Im}(T)$-এর অন্য সব বিন্দুতেও ঠিক একটা করেই তীর আসতে বাধ্য। অর্থাৎ T-টা injective হতে বাধ্য। এর উল্টোটাও অবশ্যই সত্যি--যদি T injective হয় তবে $\vec{0}_W$-তে খালি একটাই তীর আসতে পারে, সুতরাং $\ker(T)$ হবে একটা singleton set.
আমরা জানি যে $T(\vec{0}_V)=\vec{0}_W$ হবেই। সুতরাং $\vec{0}_V$ থেকে যে তীরটা বেরোবে সেটা অবশ্যই $\vec{0}_W$-তে পৌঁছবে। সুতরাং $\ker(T)$-এ আর কিছু থাক বা নাই থাক $\vec{0}_V$ অন্ততঃ থাকবেই। অতএব $\ker(T)$ যদি একটা singleton set হয়, তবে $\ker(T)=\{\vec{0}_V\}$ হবেই।
এই হল $\ker(T)$-এর সঙ্গে T-এর injective হওয়ার সম্পর্ক--

$$T \text{ is injective } \iff \ker(T)=\{\vec{0}_V\}.$$

আরেকটা জিনিসও লক্ষ্য কর। $\ker(T)=\{\vec{0}_V\} \iff \text{nullity}(T)=0$. সুতরাং

$$T \text{ is injective } \iff \text{nullity}(T)=0.$$

সেটাই theorem হিসেবে লিখেছি নীচে--

THEOREM

Let V and W be two vector spaces over a field F, and let $T:V\to W$ be a linear transformation. Then the following three conditions are equivalent:

1. T is one-one
2. $\ker(T)=\{\vec{0}_V\}$
3. $\text{nullity}(T)=0$.

Proof: শেষের দুটো যে equivalent, সে তো বলাই বাহুল্য--

> We know that a subspace has 0 dimension if and only if it is trivial. So (2) $\iff$ (3).

পড়ে রইল (1)-কে (2) বা (3) কারো একটার সাথে equivalent দেখানো। আমরা (2)-এর সাথে equivalent দেখাব।

> (1) $\implies$ (2):
>
> Since T is injective, $\forall \vec{u}, \vec{v} \in V \quad \big(T(\vec{u}) = T(\vec{v}) \implies \vec{u} = \vec{v}\big)$.

অর্থাৎ যদি $\vec{u}, \vec{v} \in V$ এমন হয় যে, T ওদের একই জায়গায় নিয়ে যাচ্ছে, তবে ওরা সমান হতে বাধ্য।

> We know that $T(\vec{0}_V) = \vec{0}_W$.
>
> Take any $\vec{v} \in \ker(T)$. Then $T(\vec{v}) = \vec{0}_W$.
>
> Since $T(\vec{v}) = T(\vec{0}_V)$, hence $\vec{v} = \vec{0}_V$.
>
> So $\ker(T) = \{\vec{0}_V\}$, as required.

এবার অন্য দিকটা--

> (2) $\implies$ (1):
>
> Let $\ker(T) = \{\vec{0}_V\}$.
>
> Shall show
>
> $\forall \vec{u}, \vec{v} \in V \quad \big(T(\vec{u}) = T(\vec{v}) \implies \vec{u} = \vec{v}\big)$.
>
> Take any $\vec{u}, \vec{v} \in V$ such that $T(\vec{u}) = T(\vec{v})$.
>
> Then
>
> $$\begin{aligned} & T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v}) = \vec{0}_W \\ \implies \quad & \vec{u} - \vec{v} \in \ker(T) \\ \implies \quad & \vec{u} - \vec{v} = \vec{0}_V \quad \big[\because \ker(T) = \{\vec{0}_V\}\big] \\ \implies \quad & \vec{u} = \vec{v}, \end{aligned}$$
>
> as required.

[Q.E.D]

29.4 Matrix-এর সঙ্গে সম্পর্ক

কোনো linear transformation আর (কোনো basis-এর সাপেক্ষে) তার matrix, এই দুটো জিনিস বলতে গেলে একই ধারণার দুটো রূপ, তাই একটার ধর্ম জানলেই অন্যটার ধর্মও জানা হয়ে যায়। এই ব্যাপারটারই আরেকটা উদাহরণ দেখব এবার। একটা linear transformation দেওয়া থাকলে, সেটা injective নাকি surjective নাকি bijective, নাকি কোনোটাই নয়, সেটা বলে দেওয়া যায় (যেকোনো একজোড়া basis-এর সাপেক্ষে) তার matrix-টা থেকেই। সেটাই এবার শিখব আমরা। প্রথমে injective নিয়ে শুরু করি।

THEOREM

Let V and W be vector spaces of finite dimensions over a field F and let $T : V \to W$ be a linear transformation. Let A be the matrix of T (wrt some bases of V and W) Then T is injective if and only if A is full column rank.

Proof: প্রমাণ অতি সংক্ষিপ্ত। আগেই দেখেছি injective হওয়া মানে $\text{nullity}(T) = 0$. এদিকে $\text{nullity}(T) = \text{nullity}(A)$. তাই T-এর injective হওয়া আর $\text{nullity}(A) = 0$ হওয়া সমার্থক। যদি A-র column-সংখ্যা n হয়, তবে আমরা জানি--

$$r(A) + \text{nullity}(A) = n.$$

তার মানে এখানে $r(A) = n$ হচ্ছে, অর্থাৎ column-গুলো সবাই linearly independent. তাকেই তো full column rank বলে! **[Q.E.D]**

Injective-এর জন্য theorem-টা দেখলে। এ থেকে surjective-এর জন্য নীচের theorem-টা আন্দাজ করে ফেলা কঠিন নয়--

THEOREM

Let V and W be vector spaces of finite dimensions over a field F and let $T : V \to W$ be a linear transformation. Let A be the matrix of T (wrt some bases of V and W). Then T is surjective if and only if A is full row rank.

Exercise 47: এর প্রমাণটা আরো ছোটো। তোমার জন্যেই রেখে দিলাম। আগের প্রমাণে kernel নিয়ে কাজ করেছিলাম। এবার image নিয়ে করতে হবে। ■

একটা function যদি একই সঙ্গে injective এবং surjective হয়, তবেই তাকে bijective বলে। সুতরাং আগের দুটো theorem মিলিয়ে আমরা বুঝতে পারছি যে, একটা linear transformation একমাত্র তখনই bijective হবে, যখন তার matrix হবে একাধারে full column rank এবং full row rank, অর্থাৎ যখন matrix-টা nonsingular হবে। আমরা আগেই দেখেছি যে, এক্ষেত্রে inverse-টা ফের একটা linear transformation হবে। এই linear transformation-টার matrix-টা হবে শুরুর matrix-টার inverse.

THEOREM

Let V and W be vector spaces of finite dimensions over a field F and let $T : V \to W$ be a linear transformation. Let A be the matrix of T (wrt some bases of V and W). Then T is bijective if and only if A is nonsingular. Also then the matrix of T^{-1} wrt the same bases is A^{-1}.

Proof: Theorem-টার প্রথম অংশ তো আগের theorem দুটো মেলালেই আসে। আমরা খালি শেষের অংশটা প্রমাণ করব, মানে inverse-এর জায়গাটা।

> We first fix some basis $\mathcal{B}_V$ of V, and some basis $\mathcal{B}_W$ of W.
> Let T be a bijective linear transformation having matrix A wrt $\mathcal{B}_V, \mathcal{B}_W$.
> Then A is nonsingular.

এইটা এসেছে theorem-এর প্রথম অংশটা থেকে।

> Let $B = A^{-1}$.
> Let $S : W \to V$ be the linear transformation represented by B wrt the bases $\mathcal{B}_W, \mathcal{B}_V$.

লক্ষ্য করো, এবার $\mathcal{B}_W$-কে আগে লিখেছি।

> Then clearly, $S \circ T : V \to V$ is a linear transformation represented by BA wrt $\mathcal{B}_V$

কারণ linear transformation-দের composition নিলে তাদের matrix-গুলো গুণ হয়ে যায়। যেহেতু এখানে domain আর codomain একই, তাই একটা basis-এই কাজ চলে যাচ্ছে।

> But $BA = I$. Hence $S \circ T$ must be the identity linear transformation on V.
> Similarly, $AB = I$, and hence $T \circ S$ must be the identity linear transformation on W.
> Thus S is the inverse of T. Since S is a linear transformation, the proof is complete.

[Q.E.D]

DAY 30
Spanning আর linearly independent

ধরো একটা linear transformation আছে, $T : V \to W$. এই T-টা V আর W-এর মধ্যে একটা যোগসূত্র রচনা করছে। যদি কিছু $\vec{v}_1, ..., \vec{v}_n \in V$ দেওয়া থাকে, তবে তা থেকে তুমি W-এর মধ্যে $T(\vec{v}_1), ..., T(\vec{v}_n)$ পাবে। অনুমান করা শক্ত নয় যে, $\{\vec{v}_1, ..., \vec{v}_n\}$-এর বিভিন্ন ধর্মের সঙ্গে $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$-এর বিভিন্ন ধর্মের যোগ থাকবে। সেই যোগটা আলোচনা করাই আজকে আমাদের উদ্দেশ্য। শুরুতেই সংক্ষেপে বলে নিই আমরা কী দেখাতে চলেছি--

- যদি $\{\vec{v}_1, ..., \vec{v}_n\}$ হয় V-এর একটা spanning set, তবে $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ হবে $\text{Im}(T)$-এর একটা spanning set. (সাবধান, পুরো W-এর spanning set বলিনি কিন্তু!)
- এর উল্টো দিকটা ঠিক নাও হতে পারে, মানে $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ যদি $\text{Im}(T)$-এর spanning set হয়ও, তাও $\{\vec{v}_1, ..., \vec{v}_n\}$ কিন্তু V-এর একটা spanning set নাও হতে পারে। এরকম counterexample পাওয়া খুবই সহজ। একটু পরেই একটা দেব। তার আগে চেষ্টা করে দেখতে পারো, নিজেই ভেবে বার করতে পারো কিনা!
- Linearly independent-এর বেলায় গল্পটা ঠিক উল্টো। যদি $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ একটা linearly independent set হয়, তবে $\{\vec{v}_1, ..., \vec{v}_n\}$-ও তাই হবে।

- কিন্তু $\{\vec{v}_1, ..., \vec{v}_n\}$ যদি linearly independent হয়, তবে $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ তা নাও হতে পারে। এর counterexample-ও খুব সহজ। ভেবে দেখো!

এর মধ্যে দ্বিতীয় আর চতুর্থ পয়েন্টে উল্টোদিকগুলো কাজ করছিল না। যদি T-টা injective হয় (মানে one-one হয়), তবে সেগুলোও কাজ করবে।

T যখন injective, তখন--

- যদি $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ হয় $\mathrm{Im}(T)$-এর একটা spanning set, তবে $\{\vec{v}_1, ..., \vec{v}_n\}$-ও হবে V-এর একটা spanning set. আর $\{\vec{v}_1, ..., \vec{v}_n\}$-টা linearly independent হলে পরে $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$-ও তাই হবে।

T-টা injective হলে কী সুবিধা হবে বললাম। স্বভাবতঃই প্রশ্ন ওঠে T-টা surjective (মানে onto) হলে কী বাড়তি সুবিধা হতে পারে। উত্তরটা অতি সহজ এবং সংক্ষিপ্ত--এতক্ষণ যেখানে যেখানে $\mathrm{Im}(T)$ লিখছিলাম, সে সব জায়গায় এবার W লেখা যাবে, কারণ তখন তো $\mathrm{Im}(T) = W$.
এই হল আজকের পুরো পড়ার মূল নির্যাস। এবার সবকিছু গুছিয়ে লিখে প্রমাণ করব এবং কিন্তু প্রয়োগ দেখব।

30.1 Spanning

THEOREM

Let $T : V \to W$ be a linear transformation. Let $\vec{v}_1, ..., \vec{v}_n \in V$. If $V = \mathrm{span}\{\vec{v}_1, ..., \vec{v}_n\}$, then $\mathrm{Im}(T) = \mathrm{span}\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$.

Proof:

Let the field be F.
Shall show that
$\forall \vec{w} \in \mathrm{Im}(T) \quad \exists a_1, ..., a_n \in F \quad \vec{w} = \sum_{i=1}^n a_i T(\vec{v}_i)$.

এইটা এসেছে spanning set-এর সংজ্ঞা থেকে, যেকোনো $\vec{w} \in \mathrm{Im}(T)$-কেই $T(\vec{v}_i)$-দের linear combination হিসেবে লেখা যাবে।

$\forall \vec{w}$ Take any $\vec{w} \in \mathrm{Im}(T)$.
Then $\vec{w} = T(\vec{v})$ for some $\vec{v} \in V$.

এটা আসছে image-এর সংজ্ঞা থেকে।

Since $V = \mathrm{span}\{\vec{v}_1, ..., \vec{v}_n\}$,
hence $\exists a_1, ..., a_n \in F \quad \vec{v} = \sum_{i=1}^n a_i \vec{v}_i$.
$\exists a_i$ Choose these a_i's.

Then

$$\begin{aligned} \vec{w} &= T(\vec{v}) \\ &= T\left(\sum_{i=1}^{n} a_i \vec{v}_i\right) \\ &= \sum_{i=1}^{n} a_i T(\vec{v}_i) \quad [\because T \text{ is linear}], \end{aligned}$$

as required.

[Q.E.D]

আগেই বলেছিলাম যে, এর উল্টো দিকটা সর্বদা খাটে না, মানে এমন $T : V \to W$ এবং $\vec{v}_1, ..., \vec{v}_n \in V$ সম্ভব, যাতে $\text{span}\{T(\vec{v}_1), ..., T(\vec{v}_n)\} = \text{Im}(T)$ হলেও $\text{span}\{\vec{v}_1, ..., \vec{v}_n\} \neq V$ হয়। এরকম একটা counterexample ভাবতে বলেছিলাম আগেই। ভেবে দেখার এই শেষ সুযোগ, কারণ আমরা এক্ষুণি একটা counterexample দিতে চলেছি।

Example 39: ধরো $T : \mathbb{R}^2 \to \mathbb{R}$ হল এইরকম, $T(x, y) \equiv 0$. তবে এমন একটা $n \in \mathbb{N}$ এবং $\vec{v}_1, ..., \vec{v}_n \in \mathbb{R}^2$ দাও, যাতে $\text{span}\{T(\vec{v}_1), ..., T(\vec{v}_n)\} = \text{Im}(T)$ হয়, অথচ $\text{span}\{\vec{v}_1, ..., \vec{v}_n\}$ মোটেই $\mathbb{R}^2$ না হয়।

SOLUTION: এখানে $\text{Im}(T) = \{0\}$. আমরা $n = 1$ এবং $\vec{v}_1 = (1, 1)$ নিতে পারি। তবে $T(\vec{v}_1) = 0$. যেহেতু 0-কে 1×0 হিসেবে ভাবাই যায়, তাই বলতে পারি যে, $\text{Im}(T)$-এর সদস্যদের $T(\vec{v}_1)$-এর linear combination হিসেবে লেখা গেল। মানছি, ব্যাপারটা খুব বোকা-বোকা শোনাচ্ছে, কিন্তু এর মধ্যে ভুল নেই কিছু। তাই $\text{Im}(T) = \text{span}\{T(\vec{v}_1)\}$. কিন্তু $\text{span}\{\vec{v}_1\}$ তো খালি $(1, 1)$ বরাবর লাইনটা, সেটা পুরো $\mathbb{R}^2$ নয় মোটেই। ■

এই counterexample-টায় যদি তোমার মন না ওঠে, তবে নীচের অংকে আরেকটা counterexample বানাতে দিয়েছি, যার মধ্যে বোকা-বোকা ব্যাপার কিছু নেই।

Exercise 48: ধরো $T : \mathbb{R}^3 \to \mathbb{R}^2$ হল $T(x, y, z) = (x, y)$. এবার এমন $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^3$ দাও, যাতে $\text{span}\{T(\vec{v}_1), T(\vec{v}_2)\} = \text{Im}(T)$ হয়, অথচ $\text{span}\{\vec{v}_1, \vec{v}_2\}$ যেন $\mathbb{R}^3$ না হয়। ■

এতক্ষণ আমরা spanning নিয়ে আলোচনা করছিলাম। এবার linear independence-এর দিকে চোখ ফেরানো যাক।

30.2 Linear independence

THEOREM

Let $T : V \to W$ be a linear transformation. Let $\vec{v}_1, ..., \vec{v}_n \in V$. If $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ is linearly independent, then so is $\{\vec{v}_1, ..., \vec{v}_n\}$.

Proof: Linearly independent দেখানোর সেই পরিচিত কায়দাই লাগাব, যেকোনো একটা linear combination-কে $\vec{0}$ ধরে coefficient-গুলোকে আলাদা করে শূন্য দেখানো--

> Let $a_1, ..., a_n \in F$ be such that
> $$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}_V. \quad (*)$$
> Shall show $a_1 = \cdots = a_n = 0$.
> Applying T to both sides of (*),
> $$T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) = T(\vec{0}_V),$$
> or, by linearity of T,
> $$a_1T(\vec{v}_1) + \cdots + a_nT(\vec{v}_n) = \vec{0}_W.$$
> But $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ is linearly independent.
> So $a_1 = \cdots = a_n = 0$, as required.

[Q.E.D]

এই theorem-এর উল্টো দিকটা যে সত্য নয়, সেটা আমরা আগেই বলেছিলাম। অর্থাৎ যদি $\{\vec{v}_1, ..., \vec{v}_n\}$ একটা linearly independent হয়ও, তাও $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ কিন্তু linearly independent নাও হতে পারে। নীচের অংকে এরকম একটা counterexample চেয়েছে। চেষ্টা করে দ্যাখো। না পারলে তো উত্তর দেখে নেওয়া আছেই।

Exercise 49: ধরো $T : \mathbb{R}^2 \to \mathbb{R}$ হল $T(x, y) = x$. এবার $\mathbb{R}^2$-র মধ্যে এমন linearly independent $\vec{v}_1, \vec{v}_2$ দাও, যাতে $\{T(\vec{v}_1), T(\vec{v}_2)\}$-টা $\mathbb{R}$-এর মধ্যে linearly independent না হয়। ■

এই যে আমরা দুই ধরণের counterexample দেখলাম, spanning-এর জন্য আর linearly independent-এর জন্য, এই দুই ধরণের counterexample-ই বন্ধ হয়ে যেত, যদি T-টা injective হত। এবার সেটাই আলোচনা করব।

30.3 Injective-এর জন্য

যদি T-টা injective হয়, তবে linearly independent set-রা T লাগানোর পরও linearly independent-ই থাকবে--

> THEOREM
>
> Let $T : V \to W$ be any injective linear transformation. Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be linearly independent in V. Then $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ must be linearly independent in W.

Proof: প্রমাণ জলের মত সহজ।

> Let $a_1T(\vec{v}_1) + \cdots + a_nT(\vec{v}_n) = \vec{0}_W$ for some scalars $a_1, ..., a_n$.
> Shall show $a_1 = \cdots = a_n = 0$.
> By linearity of T
> $$T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) = \vec{0}_W.$$
> Also $T(\vec{0}_V) = \vec{0}_W$.

> Since T is injective, so $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}_V$.

এটা কেন হল, বুঝলে তো? T-এর প্রভাবে $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$ এবং $\vec{0}_V$ দুজনেই একই জায়গায় যাচ্ছে। যেহেতু T হল injective, তাই ওরা দুজনে সমান না হয়ে যায় না।

> Since $\vec{v}_1, ..., \vec{v}_n$ are linearly independent, hence $a_1 = \cdots = a_n = 0$, as required.

[Q.E.D]

একটা অংক কষে theorem-টা হজম করে নিই--

Exercise 50: A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^4$ is defined by

$$T(x_1, x_2, x_3) = (x_2 + x_3,\ \ x_3 + x_1,\ \ x_1 + x_2, x_1 + x_2 + x_3), \quad (x_1, x_2, x_3) \in \mathbb{R}^3.$$

Find $\ker(T)$. What conclusion can you draw regarding the linear dependence/independence of the image set of the set of vectors

$$\{(1, 0, 0),\ \ (0, 1, 0),\ \ (0, 0, 1)\}?$$

HINT:
এখানে কিন্তু injective বলে দেয়নি। ■

আমরা দেখেছি যে, $\text{span}\{T(\vec{v}_1), ..., T(\vec{v}_n)\} = \text{Im}(T)$ হলেই আমরা দাবী করতে পারি না যে, $\text{span}\{\vec{v}_1, ..., \vec{v}_n\} = V$ হবে। কিন্তু T-টা injective হলে সে দাবিটা করা যায়। সেটাই প্রমাণ করব এবার।

THEOREM

Let $T : V \to W$ be any injective linear transformation. Let $\vec{v}_1, ..., \vec{v}_n \in V$. If $\text{span}\{T(\vec{v}_1), ..., T(\vec{v}_n)\} = \text{Im}(T)$, then $\text{span}\{\vec{v}_1, ..., \vec{v}_n\} = V$.

Proof:

> To show
> $\forall \vec{v} \in V \quad \exists a_1, ..., a_n \in F \quad \vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$.
>
> [$\forall \vec{v}$] Take any $\vec{v} \in V$.
> Then $T(\vec{v}) \in \text{Im}(T)$.
> Since $\text{Im}(T) = \text{span}\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$,
> hence $T(\vec{v}) = a_1T(\vec{v}_1) + \cdots + a_nT(\vec{v}_n)$ for some $a_1, ..., a_n \in F$.
>
> [$\exists a_i$] Choose these a_i's.
>
> We have $T(\vec{v}) = T(a_1\vec{v}_1) + \cdots + T(a_n\vec{v}_n)$.
> Since T is injective, this means $\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$, as required.

∎

[Q.E.D]

30.4 ফলশ্রুতি

আমরা এই যে linearly independent set আর spanning set-দের উপরে linear transformation-এর প্রভাবের কথা শিখলাম, সে সব মিলিয়েজুলিয়ে আরো নানারকম জিনিস বানানো যায়। তার কিছু নমুনা দেখব এবার।

Exercise 51: If $T : V \to W$ is a linear transformation, where V, W are finite dimensional, and U is a subspace of V, then show that $T(U)$ is a subspace of W and $\dim(T(U)) \leq \dim(U)$.
HINT:
এখানে $T(U)$ মানে হল $\{T(\vec{u}) \; : \; \vec{u} \in U\}$. এটা যে W-এর একটা subspace, সেটা সহজেই দেখাতে পারা উচিত। এবার $T(U)$-এর একটা basis নাও, এবং সেখান থেকে পিছন দিকে যাওয়ার চেষ্টা করো। ■

Exercise 52: আগের অংকে যদি বাড়তি বলে দেওয়া থাকত যে, T হল one-one, তবে কি $\dim(T(U)) = \dim(U)$ বলা যেত? ■

Generating set-এর উপর একটা linear transformation লাগালে যে image-এর generating set পাওয়া যায়, সে তো আগেই বলেছি, তার জন্য linear transformation-টার উপর কোনো বাড়তি শর্তও লাগে না। আরও দেখেছি যে, one-one হলে linear independence-ও অক্ষুণ্ণ থাকে। আমরা জানি যে generating set যদি linearly independent হয় তবে তাকেই বলে basis. অতএব নীচের অংকটা মোটেই অপ্রত্যাশিত নয়।

Example 40: Let V and W be vector spaces over a field F, and $T : V \to W$ be a linear transformation and $\{\vec{v}_1, ..., \vec{v}_n\}$ be a basis of V and $\ker(T) = \{\vec{0}_V\}$. Prove that $\{T(\vec{v}_1), T(\vec{v}_2), ..., T(\vec{v}_n)\}$ is a basis of $\text{Im}(T)$.
SOLUTION: এখানে $\ker(T) = \{\vec{0}_V\}$ বলে দিয়েছে। তার মানে ঘুরিয়ে বলে দেওয়া হল যে, T হচ্ছে injective. প্রমাণটাকে দুই ধাপে ভেঙে করব।

Step 1: Shall show that $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ is a linearly independent set.
Let $a_1, ..., a_n \in F$ be such that

$$a_1 T(\vec{v}_1) + \cdots + a_n T(\vec{v}_n) = \vec{0}_W.$$

Shall show that $a_1 = \cdots = a_n = 0$.
By linearity of T,

$$T(a_1 \vec{v}_1 + \cdots + a_n \vec{v}_n) = \vec{0}_W,$$

or

$$a_1 \vec{v}_1 + \cdots + a_n \vec{v}_n \in \ker(T).$$

But $\ker(T) = \{ \vec{0}_V\}$, and so

$$a_1 \vec{v}_1 + \cdots + a_n \vec{v}_n = \vec{0}_V.$$

But $\{\vec{v}_1, ..., \vec{v}_n\}$ is a linearly independent set in V.
So this implies $a_1 = \cdots = a_n = 0$, as required.

এই যুক্তিটা তোমার অপরিচিত হওয়ার কথা নয়, কারণ একটু আগেই আমরা একই যুক্তি ব্যবহার করেছি। এবার দ্বিতীয় ধাপে spanning দেখাব। এখানে যে যুক্তিটা লাগাব, সেটাও আমরা আগেই দেখেছি।

Step 2: Shall show that $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ generates $\text{Im}(T)$, ie,

$\forall \vec{w} \in \text{Im}(T) \quad \exists b_1, ..., b_n \in F \quad \vec{w} = \sum_1^n b_i T(\vec{v}_i)$.

$\boxed{\forall \vec{w}}$ Take any $\vec{w} \in \text{Im}(T)$.

Then $\vec{w} = T(\vec{v})$ for some $\vec{v} \in V$.

So $\vec{v} = \sum_1^n b_i \vec{v}_i$ for some $b_1, ..., b_n \in F$.

$\boxed{\exists b_i}$ Choose these b_i's.

Then, by linearity of T,

$$\begin{aligned} \vec{w} &= T(\vec{v}) \\ &= T\left(\sum_1^n b_i \vec{v} i\right) \\ &= \sum_1^n b_i T(\vec{v}_i), \end{aligned}$$

as required.

■

এবার একটা সামান্য ঠকানো অংক দিই।

Exercise 53: Let V and W be two vector spaces over a field F, and let $T : V \to W$ be an injective linear transformation. Then is the following statement true?

"The image of a basis of V is a basis of W."

If it is true, then prove it. Otherwise, provide a counterexample.
HINT:
এখানে image of basis মানেটা বুঝে নাও। যদি $\{\vec{v}_1, ..., \vec{v}_n\}$ হয় V-এর একটা basis, তবে তার image হল $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$. ■

Exercise 54: V and W are two finite dimensional vector spaces over a field. The mapping $T : V \to W$ is linear and bijective. Let if $\{\vec{v}_1, , ..., \vec{v}_n\} \subseteq V$. If $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ is a basis of W, must $\{\vec{v}_1, ..., \vec{v}_n\}$ be a basis of V? ■

Exercise 55: Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T(x_1, x_2, x_3) = (2x_1 + x_2 - x_3,\ \ x_2 + 4x_3,\ \ x_1 - x_2 + 3x_3).$$

Show that T maps a basis of $\mathbb{R}^3$ onto a basis of $\mathbb{R}^3$. ■

DAY 31
Linear transformation-দের কিছু ধর্ম

এবার আমরা linear transformation-দের চারটে ধর্ম শিখব। প্রথম তিনটে সব ক্ষেত্রেই কাজ করে, শেষেরটা কেবল finite dimension-এর জন্য।

31.1 $L(V, W)$

আমরা এবার যে আলোচনাটা করতে যাচ্ছি, সেটা finite বা infinite dimensional দুই ক্ষেত্রেই সমানভাবে প্রযোজ্য। কিন্তু প্রথমে finite dimensional কেসটা ভেবে নিলে বুঝতে সুবিধা হবে। ধরো $T : \mathbb{R}^n \to \mathbb{R}^m$ হল একটা linear transformation. তাহলে T-কে আমরা একটা $m \times n$ সাইজের matrix হিসেবে ভাবতে পারি। এবার লক্ষ কর যে vector এবং matrix দুটো জিনিসকেই বস্তুতঃ একগুচ্ছ সংখ্যা দিয়ে লেখা যাচ্ছে, তফাৎ খালি এই যে vector-এর বেলায় সংখ্যাগুলোকে একলাইনে লিখছি, আর matrix-এর বেলায় rectangle-এর মত করে। সুতরাং বলা যেতেই পারে যে, matrix-রাও আসলে ভেতরে ভেতরে vector-ই, যেমন $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$-কে $(1, 2, 3, 4)$ ভাবলেই হল। দুটো matrix-এর যোগ সেক্ষেত্রে vector-এর যোগে পরিণত হয়। কোনো matrix-কে কোনো scalar দিয়ে গুণ করলে সেই vector-টাও ওই scalar দিয়ে গুণ হয়ে যায়। সুতরাং এতে নিশ্চয়ই সন্দেহ নেই যে, যাবতীয় $m \times n$ সাইজের (real) matrix-দের সেট আসলে $\mathbb{R}^{mn}$ নামক vector space-টাই, খালি vector-গুলোকে একলাইনে না লিখে, rectangle-এর আকারে লেখা হয়েছে, এই যা তফাৎ। যেহেতু finite dimensional ক্ষেত্রে matrix আর linear transformation-কে একই জিনিসের রূপভেদ বলে মনে করা যায়, তাই বলতে পারি যে, $\mathbb{R}^n$ থেকে $\mathbb{R}^m$-এ যাবতীয় linear transformation-রা মিলে একটা vector space সৃষ্টি করে।
মজার কথা হল যে একই কথা বলা চলে infinite dimensional ক্ষেত্রেও! সেটাই বলেছে নীচের theorem-এ।

THEOREM

Let V, W be two vector spaces both defined on the same field, F. Let $L(V, W)$ be the set of all linear transformations from V to W. We define addition and scalar multiplication on $L(V, W)$ as follows: For any $T, S \in L(V, W)$ and any $\alpha \in F$,

$$\begin{aligned}(T + S)(\vec{v}) &= T(\vec{v}) + S(\vec{v}),\\ (\alpha T)(\vec{v}) &= \alpha T(\vec{v})\end{aligned}$$

for $\vec{v} \in V$. Then $L(V, W)$ is a vector space over F under these operations.

যেভাবে আমরা addition আর scalar multiplication-এর সংজ্ঞা দিলাম, তাকে অনেকসময়ে বলে **pointwise definition**, যেমন $T + S$-এর সংজ্ঞায় আমরা প্রথমে একটা $\vec{v} \in V$ নিলাম (সেটা যেন একটা point), তারপর সেই point-এ T আর S-এর value বার করে তাদের যোগ করে দিলাম।

Proof: প্রমাণটা বিরক্তিকর রকমের লম্বা, কিন্তু এর মধ্যে জটিলতা কিছু নেই। Vector space-এর সংজ্ঞায় যে একরাশ শর্ত আছে, সেগুলো একে একে পরীক্ষা করে যেতে হবে।
প্রথমে আমরা দেখাব "closure under +", অর্থাৎ $T, S \in L(V, W)$ হলে $T + S$-ও ফের $L(V, W)$-এর মধ্যেই থাকবে, এটা দেখানো। $T + S \in L(V, W)$ হওয়া মানে হল যেকোনো $\vec{v}, \vec{u} \in V$ আর যেকোনো $\alpha, \beta \in F$-এর জন্যই $(T + S)(\alpha\vec{v} + \beta\vec{u}) = \alpha(T + S)(\vec{v}) + \beta(T + S)(\vec{u})$ হওয়া। সেই কথাটা লিখে শুরু করি।

Closure under $+$: We shall show

$$\forall T, S \in L(V, W) \quad \forall \vec{u}, \vec{v} \in V \quad \forall \alpha, \beta \in F \quad (T + S)(\alpha\vec{u} + \beta\vec{v}) = \alpha(T + S)(\vec{u}) + \beta(T + S)(\vec{v}).$$

এবার এই ঘটোৎকচটাকে প্রমাণ করব।

$\forall T, S$ Take any $T, S \in L(V, W)$.

$\forall \vec{u}, \vec{v}$ Take any $\vec{u}, \vec{v} \in V$.

$\forall \alpha, \beta$ Take any $\alpha, \beta \in F$.

Then

$$\begin{aligned}(T + S)(\alpha\vec{u} + \beta\vec{v}) &= T(\alpha\vec{u} + \beta\vec{v}) + S(\alpha\vec{u} + \beta\vec{v}) \quad [\text{ by defn of } T + S] \\ &= \alpha T(\vec{u}) + \beta T(\vec{v}) + \alpha S(\vec{u}) + \beta S(\vec{v}) \quad [\because T, S \text{ are linear}] \\ &= \alpha(T + S)(\vec{u}) + \beta(T + S)(\vec{v}) \quad [\text{ by defn of } T + S],\end{aligned}$$

as required.

এবার একইভাবে এগোতে হবে scalar multiplication-এর closure নিয়ে--

Closure under scalar multiplication: To show:

$$\forall T \in L(V, W) \quad \forall \vec{u}, \vec{v} \in V \quad \forall a, \alpha, \beta \in F \quad (aT)(\alpha\vec{u} + \beta\vec{v}) = \alpha(aT)(\vec{u}) + \beta(aT)(\vec{v}).$$

এর প্রমাণটা তোমার জন্য রেখে দিলাম।

Associativity of $+$: To show

$$\forall T, S, U \in L(V, W) \quad \forall \vec{v} \in V \quad ((T + S) + U)(\vec{v}) = (T + (S + U))(\vec{v}).$$

$\forall T, S, U$ Take any $T, S, U \in L(V, W)$.

$\forall \vec{v}$ Take any $\vec{v} \in V$.

Then

$$
\begin{aligned}
((T+S)+U)(\vec{v}) &= (T+S)(\vec{v})+U(\vec{v}) \quad [\text{ by defn of } +] \\
&= \big(T(\vec{v})+S(\vec{v})\big)+U(\vec{v}) \quad [\text{ by defn of } +] \\
&= T(\vec{v})+\big(S(\vec{v})+U(\vec{v})\big) \quad [\text{ by asso of } + \text{ in } W]
\end{aligned}
$$

এখানটা গোলমেলে লাগতে পারে। আমরা তো '+'-এর associativity-ই, প্রমাণ করছি। তবে '+'-এর associativity ব্যবহার করছি কী করে? আসলে আমরা দেখাচ্ছি linear transformation-দের '+'-এর associativity, আর ব্যবহার করছি W-এর '+'-এর associativity.

$$
\begin{aligned}
&= T(\vec{v})+(S+U)(\vec{v}) \quad [\text{ by defn of } +] \\
&= (T+(S+U))(\vec{v}) \quad [\text{ by defn of } +],
\end{aligned}
$$

as required.

একইরকম পথে commutativity-ও প্রমাণ হয়ে যাবে।

Commutativity of $+$:

To show

$\forall T,S\in L(V,W) \quad \forall \vec{v}\in V \quad (T+S)(\vec{v})=(S+T)(\vec{v}).$

ধাপগুলো তোমার জন্যই রেখে দিলাম। এরপর দেখাতে হবে additive identity-র অস্তিত্ব। আন্দাজ করা যাচ্ছে যে, additive identity হবে সেই linear transformation-টা যেটা সব কিছুকেই $\vec{0}_W$-তে নিয়ে ফেলে।

Existence of additive identity:

To show:

$\exists Z\in L(V,W) \quad \forall T\in L(V,W) \quad \forall \vec{v}\in V \quad (T+Z)(\vec{v})=T(\vec{v}).$

Consider the function $Z:V\to W$ defined as

$$\forall \vec{v}\in V \quad Z(\vec{v})=\vec{0}_W.$$

This is a linear transformation.

⟦ Because:

$\forall a,b\in F \quad \forall \vec{v}_1,\vec{v}_2\in V$

$$
\begin{aligned}
Z(a\vec{v}_1+b\vec{v}_2) &= \vec{0}_W \\
&= a\vec{0}_W+b\vec{0}_W \\
&= aZ(\vec{v}_1)+bZ(\vec{v}_2).
\end{aligned}
$$

⟧

$\boxed{\exists Z}$ Choose this $Z\in L(V,W)$.

$\forall T$ Take any $T \in L(V, W)$.

$\forall \vec{v}$ Take any $\vec{v} \in V$.

Then $(T+Z)(\vec{v}) = T(\vec{v}) + Z(\vec{v}) = T(\vec{v}) + \vec{0}_W = T(\vec{v})$, as required.

আশা করি বুঝতে পারছ যে, প্রমাণের প্রতিটা ধাপেরই আসলে একই যুক্তি, $L(V, W)$-এর যে ধর্মটা দেখাতে চাও, V আর W-এর সেই ধর্মটা ব্যবহার করলেই প্রমাণটা হয়ে যায়। অবশ্য কথাটা বলা যতটা সহজ, জীবনে প্রথমবার লেখার সময়ে ততটা সহজ নয় মোটেই। তাই বাকি ধাপগুলোর প্রতিক্ষেত্রে কী দেখাতে হবে, সেটুকু বলে ছেড়ে দিচ্ছি, তোমার কাজ হবে "মামুলী" যুক্তিগুলো বসিয়ে শূন্যস্থান পূরণ করা।

বাজনার বোল মুখে বলা সহজ, হাতে আনা শক্ত।

--শ্রীরামকৃষ্ণ

<u>Existence of additive inverse</u>: To show:

$\forall T \in L(V, W) \quad \exists S \in L(V, W) \quad \forall \vec{v} \in V \quad (T+S)(\vec{v}) = Z(\vec{v})$.

এখানে Z-টা আগের ধাপ থেকে নিয়েছি। এই S-কেই বলব $-T$. এইটা দেখানো হয়ে গেলে দুরকমের distributivity দেখাও--

<u>Distributivity</u>: To show:

$\forall T, S \in L(V, W) \quad \forall \alpha \in F \quad \forall \vec{v} \in V \quad (\alpha(T+S))(\vec{v}) = (\alpha T + \alpha S)(\vec{v})$.

$\forall T \in L(V, W) \quad \forall \alpha, \beta \in F \quad \forall \vec{v} \in V \quad ((\alpha+\beta)T)(\vec{v}) = (\alpha T + \beta T)(\vec{v})$.

তারপরের ধর্মটা দেখতে অনেকটা associativity-র মত--

To show:

$\forall T \in L(V, W) \quad \forall \alpha, \beta \in F \quad \forall \vec{v} \in V \quad ((\alpha\beta)T)(\vec{v}) = (\alpha(\beta T))(\vec{v})$.

সবশেষে খুব সহজ একটা জিনিস--

To show:

$\forall T \in L(V, W) \quad \forall \vec{v} \in V \quad (1T)(\vec{v}) = T(\vec{v})$.

এখানে 1 হল F-এর multiplicative identity. **[Q.E.D]**

31.2 Ring

আবার একটু finite dimensional কেসটা চিন্তা করে নিই, মানে linear transformation-এর জায়গায় matrix কল্পনা করে। আমরা matrix-এর যোগ-বিয়োগ আর scalar multiplication ব্যবহার করে এক্ষুণি একটা vector space বানালাম। কিন্তু matrix-এর আরেকটা গুরুত্বপূর্ণ operation আছে, সেটা হল দুটো matrix-কে গুণ করা। এবার সেটা নিয়ে কাজ করব। Matrix multiplication-এর একটা সমস্যা হল এর ফলে সাইজ বদলে যায়, যোগ-বিয়োগের বেলায় সে ঝামেলা ছিল না। সে ঝামেলা এড়াতে আমরা খালি square matrix নিয়ে কাজ করব, তাহলে এরকম দুটো matrix-কে সর্বদাই গুণ করা

যাবে এবং গুণফলও একই সাইজের একটা matrix হবে। এতে নিশ্চয়ই কোনো সন্দেহ নেই যে, matrix multiplication হল associative. আর সেটা যোগের উপর distribute-ও করে। সুতরাং যাবতীয় $n \times n$ সাইজের matrix-রা মিলে একটা ring হচ্ছে।
এ তো গেল finite dimension নিয়ে আলোচনা। Infinite dimensional ক্ষেত্রেও কি বলা যাবে যে, $L(V, V)$ একটা ring হবে? লক্ষ্য কর যে, এখানে W-এর জায়গাতেও V নিয়েছি, ঠিক যেমন matrix-এর বেলায় খালি square matrix নিয়ে কাজ করছিলাম। উত্তর হল, হ্যাঁ, এখানেও ring হবে। এটা প্রমাণ করা এতই সহজ যে তোমার জন্যই নীচের অংকটা রেখে দিলাম।

Exercise 56: Let V be any vector space. Show that $L(V, V)$ is a ring under usual addition and composition. ■

31.3 $GL(n)$

আমরা যোগ-বিয়োগ-গুণ দেখলাম। এবার তাহলে ভাগ করার দিকে চোখ ফেরাই। প্রথমে যথারীতি finite dimension নিয়ে কাজ করব, যাতে matrix ব্যবহার করে চিন্তা করা যায়। কোনো matrix দিয়ে ভাগ করা মানে যে তার inverse দিয়ে গুণ করা সেটা নিশ্চয়ই আর বলে দিতে হবে না। Inverse-এর প্রশ্ন যখন উঠল, অতএব বুঝতেই পারছ যে, এখানেও খালি square matrix-দের নিয়ে কাজ করতে হবে। খালি তাই নয়, matrix-গুলোকে nonsingular-ও হতে হবে। যদি তা হয়, তবে এটা বোঝা কঠিন নয় যে--

> যাবতীয় $n \times n$ সাইজের nonsingular matrix-রা মিলে একটা group হয়। এখানে group operation-টা হল matrix multiplication.

এই group-টার একটা নাম আছে,**general linear group** of order n বা সংক্ষেপে $GL(n)$.
আগের আলোচনার অভিজ্ঞতা থেকে আন্দাজ করতে পারছ যে, infinite dimensional ক্ষেত্রেও একইরকম একটা group পাওয়া যাবে। এখানে nonsingular matrix-দের জায়গায় থাকবে invertible linear transformation-রা।
সুতরাং এবার নীচের theorem-টা প্রমাণ করা খুবই সহজ--

> **THEOREM**
>
> Let V be any vector space. Then the set of all invertible linear transformations from V to V form a group under composition.

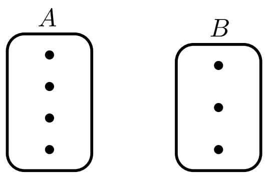
31.4 Finite set-এর মত

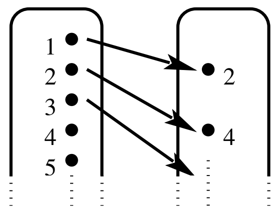
এবার আমরা একটা নতুন ধারণা শিখব। তার জন্য প্রথমে কয়েকটা সহজ অংক করি। এগুলো খালি সাধারণ function নিয়ে, এদের মধ্যে কোনো linear transformation-এর গল্প নেই।

Example 41: Fig 38-এ দুটো finite set দেখিয়েছি। তীর চিহ্ন এঁকে তোমাকে একটা function বানাতে হবে,

$f : A \to B$, যেটা injective হয়।
SOLUTION: একটু চেষ্টা করলেই বুঝবে যে কাজটা করা যাচ্ছে না, কারণ B-তে A-র চেয়ে কম বিন্দু আছে। ■

এই সমস্যাটা কিন্তু infinite set-এর ক্ষেত্রে নাও হতে পারত। যেমন ধরো, যদি $\mathbb{N}$ আর $2\mathbb{N}$ নিই ($2\mathbb{N}$ হল যাবতীয় even, মানে জোড় সংখ্যার set), তবে দেখতেই পাচ্ছ যে $2\mathbb{N}$-এর মধ্যে $\mathbb{N}$-এর চেয়ে "কম" সংখ্যা আছে (কারণ বিজোড়গুলো বাদ গেছে), অথচ $f(n) = 2n$ একটা injective function. Fig 39 দ্যাখো।





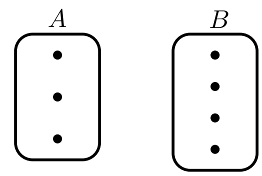


Fig 38 **Fig 39** **Fig 40**

এখন লক্ষ্য কর যে, vector space-রা প্রায়শঃই infinite set হয়। যেমন, যদি field-টা infinite হয় (যেমন $\mathbb{R}$ বা $\mathbb{C}$), তবে সব nontrivial[3] vector space-ই infinite হবে। কারণটা খুবই সোজা--যদি একটাও $\vec{v} \neq \vec{0}$ থাকে, তবে $2\vec{v}, 3\vec{v}, 4\vec{v}$, ইত্যাদি যত রাজ্যের সকলেই আলাদা আলাদা vector. অতএব finite dimensional vector space-রা finite set নাও হতে পারে। কিন্তু যদি V, W দুজনেই finite dimensional হয় আর $f : V \to W$ হয় একটা linear transformation, তবে দেখবে যে আচরণটা ঠিক finite set-এর মতই হবে। এবার সেটাই আলোচনা করব। তার জন্য দুটো জিনিস মনে করিয়ে দিই--

- এক, যেকোনো linear transformation-র জন্যই rank+nullity=dim(domain) হয়।
- দুই, কোনো linear transformation-র পক্ষে injective হওয়া আর তার nullity শূন্য হওয়া সমার্থক।

Example 42: Fig 38-এর একটা vector space সংস্করণ বানানো যাক। A ছিল একটা finite set, যার সাইজ 4.

তার জায়গায় একটা vector space নেব, V, যার dimension 4. একইভাবে B-এর ভূমিকা পালন করবে একটা vector space, W, যার dimension 3. এবার প্রশ্ন হল, V থেকে W-তে কোনো injective linear transformation আছে কিনা।

SOLUTION:

যদি $f : V \to W$ এরকম একটা linear transformation হয়, তাহলে $\text{rank}(f) + \text{nullity}(f) = 4$ হতে হবে। এদিকে injective বলেছে, তার মানে $\text{nullity}(f) = 0$. সুতরাং পেলাম $\text{rank}(f) = 4$, মানে $\dim(\text{Im}(f)) = 4$, কিন্তু সেটা তো অসম্ভব, কারণ $\text{Im}(f)$ হল W-এর subspace, আর W-এর dimension-ই তো হল $3(< 4)$.

বুড়ো তার দুটো আঙুল দিয়ে আমায় একটুখানি টিপে টিপে বলল, "আড়াই সের।"

আমি বললাম, "সে কি, পটলার ওজনই তো একুশ সের, সে আমার চাইতে দেড় বছরের ছোটো।" ■

--হ য ব র ল (সুকুমার রায়)

এরকম আরো কয়েকটা উদাহরণ দেখি, যাতে finite set-দের সঙ্গে finite dimensional vector space-দের সাদৃশ্যটা ভালো করে চোখে পড়ে।

Example 43: Fig 40-তে আবার দুটো finite set দিয়েছি। এবার তোমার কাজ হল একটা surjective $f : A \to B$

বানানো। খানিকক্ষণ চেষ্টা করলেই বুঝবে করা যাচ্ছে না। এর vector space সংস্করণটা কী হবে লিখে ফেলে প্রমাণ কর তো!

SOLUTION: যদি V আর W দুটো vector space নিই (একই field-এর উপরে), যাদের dimension যথাক্রমে 3 এবং 4, তবে এমন কোনো linear transformation পাওয়া যাবে না, $f : V \to W$, যেটা surjective হয়।

কারণ, যদি যেত, তবে $\text{rank}(f) + \text{nullity}(f) = 3$ হত। এদিকে surjective বলে $\text{rank}(f) = 4$ হবে। সুতরাং $\text{nullity}(f) = -1$ হয়ে যাচ্ছে! কিন্তু $\text{nullity}(f)$ তো negative হতে পারে না! ■

[3] Trivial vector space মানে $\{\vec{0}\}$.

Exercise 57: ধরো V আর W দুটো finite dimensional vector space. যদি V থেকে W-তে কোনো injective linear transformation থাকে, তবে নীচের কোন্ কোন্ বাক্য সঠিক হতে বাধ্য?

(A) $\dim(V) \leq \dim(W)$

(B) $\dim(V) \geq \dim(W)$

(C) $\dim(V) < \dim(W)$

(D) $\dim(V) > \dim(W)$

■

Exercise 58: ধরো V আর W দুটো finite dimensional vector space. যদি V থেকে W-তে কোনো surjective linear transformation থাকে, তবে নীচের কোন্ কোন্ বাক্য সঠিক হতে বাধ্য?

(A) $\dim(V) \leq \dim(W)$

(B) $\dim(V) \geq \dim(W)$

(C) $\dim(V) < \dim(W)$

(D) $\dim(V) > \dim(W)$

■

Exercise 59: V and W are two vector spaces over a field F. Let there be injective linear transformations from V to W and also from W to V. Show that $\dim(V) = \dim(W)$. ■

আরও একটা সামান্য অন্যরকম অংক দেখাই--

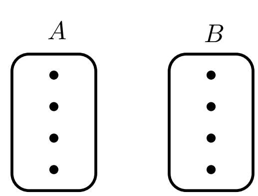
Exercise 60: এবার দুটো সমান সাইজের finite set নিয়ে শুরু করি Fig 41-এর মত। এমন একটা $f : A \to B$ দিতে পারো, যেটা injective, কিন্তু surjective নয়? পারছ না তো? এবার এর vector space সংস্করণটা লিখে প্রমাণ কর। ■

নীচের অংকটা এই একই জিনিস সামান্য অন্য ভাষায়।

Example 44: V and W are two vector spaces over a field F of same finite dimension. Prove that an injective linear transformation $T : V \to W$ is invertible.

SOLUTION:

Fig 41



> Let $\dim(V) = \dim(W) = n$.
> Let $T : V \to W$ be an injective linear transformation.
> $\because T$ is injective, $\therefore \ker(T) = \{\vec{0}_V\}$.
> Hence $\text{nullity}(T) = 0$.
> So $\text{rank}(T) = n - \text{nullity}(T) = n$.
> So $\dim(\text{Im}(T)) = n$.
> But $\text{Im}(T)$ is a subspace of W, and $\dim(W) = n$.
> Hence $\text{Im}(T) = W$.
> So T is also surjective.
> Hence T is invertible, as required.

■

নীচের অংকগুলোও সবই একই সুরের।

Exercise 61: V and W are two vector spaces of same finite dimension over a field F. Prove that a surjective linear transformation $T : V \to W$ is invertible. ■

Exercise 62: V and W are two finite dimensional vector spaces over a field F. Let $f, g : V \to W$ be two linear transformations such that f is injective and g is surjective. Show that $\dim(V) = \dim(W)$ and both f and g must be invertible. ■

DAY 32
Isomorphism

কোনো linear transformation যদি একইসঙ্গে injective এবং surjective হয় (মানে one-one এবং onto হয়), তবে তাকে বলে isomorphism of vector spaces. এটা আমরা আগেই শিখেছি। যেহেতু আমরা এখানে vector space নিয়েই আলোচনা করছি, তাই vector space কথাটা আলাদা করে বার বার বলব না, সংক্ষেপে খালি isomorphism বলব। যদি দুটো vector space-এর মধ্যে একটা isomorphism থাকে, তবে তাদেরকে বলে **isomorphic**, মানে "একইরকম গঠনবিশিষ্ট" (iso মানে "একইরকম" আর morph মানে "গঠন")। যদি V আর W দুটো পরস্পরের সঙ্গে isomorphic হয়, তবে আমরা লিখে $V \cong W$. বুঝতেই পারছ যে, $V \cong W$ আর $W \cong V$ সমার্থক। দুটো vector space যদি isomorphic হয়, তবে তাদের একটার যাবতীয় বৈশিষ্ট্য অন্যটার মধ্যেও থাকে। একটা উদাহরণ দেখি।

Example 45: P_3 হল সেই vector space, যার সদস্যদের দেখতে $a + bx + cx^2 + dx^3$-এর মত, যেখানে $a, b, c, d \in \mathbb{R}$. এখানে vector addition আর scalar multiplication যেমন polynomial-দের বেলায় হয়, সেইরকম। এবার তোমাকে একটা linear transformation দিচ্ছি $T : P_3 \to \mathbb{R}^4$ এইভাবে--

$$T(a + bx + cx^2 + dx^3) = (a, b, c, d).$$

এটা যে একটা isomorphism, তা নিয়ে আশা করি কোনো সন্দেহ[4] নেই। তাই P_3 এবং $\mathbb{R}^4$ হল isomorphic, মানে $P_3 \cong \mathbb{R}^4$. এদের দুজনের যাবতীয় ধর্মই যে একই, সে তো বোঝাই যাচ্ছে, যাকে $a + bx + cx^2 + dx^3$ লিখছিলাম, তাকেই তো মনে মনে (a, b, c, d) বলে ভাবা যায়, x-গুলো আলাদা করে না উল্লেখ করে! তার মানে P_3 আর $\mathbb{R}^4$-এর মধ্যে পার্থক্য খালি লেখার কায়দাতে, মূল কাঠামোটা একই আছে। ■

একটু আগেই বললাম যে, দুটো vector space যদি isomorphic হয়, তবে তাদের "যাবতীয় বৈশিষ্ট্য" একই হয়। এখানে কিন্তু "বৈশিষ্ট্য" বলতে খালি vector space হিসেবে বৈশিষ্ট্যগুলোর কথা বলছি। অন্য কোনো দিক দিয়ে ওদের আলাদা আলাদা বৈশিষ্ট্য থাকতেই পারে। নীচের উদাহরণটা দেখলে বুঝতে সুবিধা হবে।

Example 46: $\mathbb{C}$ এবং $\mathbb{R}^2$ দুজনেই $\mathbb{R}$-এর উপর vector space. এরা পরস্পরের সঙ্গে isomorphic. একটা isomorphism হল $T(x + iy) = (x, y)$, যেখানে $x, y \in \mathbb{R}$. কিন্তু $\mathbb{C}$-এর সদস্যদের পরস্পরের সঙ্গে গুণ করা যায়, $(a + ib)(c + id) = (ac - bd) + i(ad + bc)$. এইরকম কোনো গুণ আমরা $\mathbb{R}^2$-এর মধ্যে করি না। কিন্তু তাতে isomorphism-এর পথে বাধা হচ্ছে না, কারণ এখানে isomorphism মানে isomorphism of <u>vector spaces</u>. এই যে গুণটার কথা বললাম, সেটা vector space-এর সংজ্ঞায় ধরা নেই। ■

নীচের অংকগুলো তোমার পূর্বপরিচিত। প্রতিক্ষেত্রে একটা linear transformation দেওয়াই আছে। তোমাকে দেখাতে হবে যে, সেটা একটা isomorphism. গায়ের জোরে function-টাকে invert করার চেষ্টা করতে হবে। যেই সেটা করা যাবে, অমনি isomorphism প্রমাণ হয়ে যাবে।

Exercise 63: Prove that the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (x+y, y+z, x+z)$ is one-one and onto.[5] **(2021.CC6.3g)** ■

Exercise 64: Let T be a linear map on $\mathbb{R}^3$ defined as $T(a, b, c) = (2a, 2a - 5b, 2b + c)$. Does T^{-1} exist? If it exists, find T^{-1}.[2+3] **(2021.CC6.3f)** ■

Exercise 65: If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear mapping such that $(1, 2, 3)$, $(3, 0, 1)$ and $(0, 3, 1)$ go to $(-3, 0, -2)$, $(-5, 2, -2)$, $(4, -1, 1)$, respectively, then show that T is an isomorphism.[5] **(2021.CC6.3c)**

HINT: এই অংকটা অন্ততঃ দুভাবে করা যায়। এক, প্রথমে linear transformation-টার explicit representation-টা বার করে তারপর আগের অংকদুটোর মত এগোনো। আর দুই, এটা দেখানো যে, $\{(1, 2, 3), (3, 0, 1), (0, 3, 1)\}$ এবং $\{(-3, 0, -2), (-5, 2, -2), (4, -1, 1)\}$ দুজনেই $\mathbb{R}^3$-এর basis. যদি কোনো linear transformation একটা basis-কে আরেকটা basis-এ নিয়ে ফেলে, তবে সেটা isomorphism না হয়ে যায় না (কেন বলো তো!)। এই অংকের উত্তরটা দেখে নিলে একটা তৃতীয় কায়দা পাবে। ■

দুটো vector space যদি পরস্পরের সঙ্গে isomorphic হয়, তবে আমাদের খুব আনন্দ। কারণ সেক্ষেত্রে খালি একটার ধর্মগুলো জানলেই চলবে, অন্যটার বেলাতেও সেই একই সব ধর্ম খাটবে। বিশেষ করে যদি একটা জটিল দেখতে vector space-কে একটা সহজ দেখতে vector space-এর সঙ্গে isomorphic দেখানো যায়, তবে তো মজা আরো বেশী। এখানে লক্ষণীয় যে, isomorphic হওয়াটাই বড় কথা, isomorphism-টা ঠিক কী, সেটা বড় কথা নয়। বস্তুতঃ দুটো vector space যদি isomorphic হয়, তবে তাদের মধ্যে একাধিক isomorphism থাকতেই পারে। সন্দেহ হচ্ছে? নীচের উদাহরণটা দেখে নাও।

Example 47: আমরা এক্ষুণি বললাম যে $\mathbb{C}$ এবং $\mathbb{R}^2$ দুজনেই $\mathbb{R}$-এর উপর vector space হিসেবে isomorphic.

[4]সন্দেহ থাকলে পরীক্ষা করে দেখো!

একটা isomorphism দিয়েছিলাম এরকম, $T(x+iy) = (x, y)$, যেখানে $x, y \in \mathbb{R}$. এবার আমি আরো দুটো linear transformation দিচ্ছি $\mathbb{C}$ থেকে $\mathbb{R}^2$-তে। এরাও কি isomorphism?

$$S_1(x+iy) = (y, x) \text{ এবং } S_2(x+iy) = (x, -y)$$

SOLUTION: পরীক্ষা করেই দেখি। Linear transformation তো বলেই দিয়েছে, খালি দেখতে হবে bijection কিনা। যেমন, S_1-এর বেলায় দেখতে হবে যেকোনো $(y, x) \in \mathbb{R}^2$ দেওয়া থাকলে তা থেকে তুমি $x+iy$ বার করে দিতে পারো কিনা। অবশ্যই পারো। তাই S_1 হল একটা isomorphism. একইভাবে S_2-কে নিয়ে পরীক্ষা করে দেখো। ■

হ্যাঁ, যেটা বলছিলাম। দুটো vector space যদি isomorphic হয়, তবে সেটা আমাদের পক্ষে সুখবর। Isomorphic হবার জন্য isomorphism লাগে বটে, কিন্তু isomorphism-টা নিয়ে আমাদের মাথাব্যথা কম। তাই দুটো vector space-কে isomorphic দেখানোর সময়ে আমাদের চেষ্টা থাকে যদি কোনোভাবে কাজটা শর্টকাটে মেরে দেওয়া যায়, সরাসরি কোনো isomorphism বার না করেই। এরকম কিছু কায়দা এবার আমরা শিখব। যেমন ধরো আমরা বলেছি দুটো vector space যদি isomorphic হয়, তবে তাদের তাদের সব ধর্মই একই হয়। এবার ধরো যদি এমন একটা বিশেষ জোরালো ধর্ম পাওয়া যায়, যাতে খালি সেই একটা ধর্ম মিললেই জোর দিয়ে বলে দেওয়া যায় যে, vector space দুটো isomorphic, তবে খুব সুবিধা হবে। Finite dimensional vector space-দের বেলায় এরকম একটা জোরালো ধর্ম হল dimension. যদি একই field ০এর উপরে দুটো finite dimensional vector space-এর dimension সমান হয়, তবে তারা পরস্পরের সঙ্গে isomorphic হতে বাধ্য।

THEOREM

Two finite dimensional vector spaces V, W over a field F are isomorphic if and only if $\dim(V) = \dim(W)$.

Proof:

> If part: Let $\dim(V) = \dim(W) = n$, say.

আমরা এবার V থেকে W-তে একটা isomorphism বানিয়ে দেখাব।

> We choose two bases $\{\vec{v}_1, ..., \vec{v}_n\}$ and $\{\vec{w}_1, ..., \vec{w}_n\}$ for V and W, respectively.

আমরা সেই linear transformation-টা নেব, যেটা $\vec{v}_i$-গুলোকে $\vec{w}_i$-তে পাঠায়।

> We define a function $T : V \to W$ as the unique linear transformation such that $\forall i \quad T(\vec{v}_i) = \vec{w}_i$.

যেহেতু T-টা একটা basis-কে আরেকটা basis-এ পাঠাচ্ছে, তাই T-টা isomorphism হতে বাধ্য। সেটাই একটু বিশদ করে প্রমাণ করব এবার। প্রথমে দেখাব T-টা one-one.

> T is injective.

> ⟦ Because:
>
> Let $T(\vec{u}) = \vec{0}_W$.
> If $\vec{u} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$,
> then $T(\vec{u}) = a_1\vec{w}_1 + \cdots + a_n\vec{w}_n = \vec{0}_W$.
> So $a_1 = \cdots = a_n = 0$, as $\vec{w}_1, ..., \vec{w}_n$ are linearly independent.
> So $\ker(T)$ is trivial, and hence T is injective.
>
> ⟧

তারপর onto দেখানোর পালা--

> T is surjective.
>
> ⟦ Because:
>
> $\dim(\text{Im}(T)) = \dim(V) - \dim(\ker(T)) = \dim(V) - 0 = \dim(V) = \dim(W)$. Now $\text{Im}(T)$ is a subspace of W.
> So equality of dimension means $\text{Im}(T) = W$.
>
> ⟧
>
> Hence T is an isomorphism between V and W.
>
> So $V \cong W$, as required.

এবার উল্টো দিকটা--

> Only if part: Let $T : V \to W$ be an isomorphism.
>
> Then
>
> $$\begin{aligned} \dim(W) &= \dim(\text{Im}(T)) \quad [\because T \text{ is surjective}] \\ &= \dim(V) - \dim(\ker(T)) \\ &= \dim(V) - 0 \quad [\because T \text{ is injective}] \\ &= \dim(V), \end{aligned}$$
>
> as required.

[Q.E.D]

এই theorem-টার একটা প্রয়োগ দেখো--যদি $\mathbb{R}$-এর উপর কোনো vector space থাকে, যার dimension হল 5, তবে আর কিছুমাত্র না জেনেই চোখ বুঁজে বলে দেওয়া যায় যে, সেটা $\mathbb{R}^5$-এর সাথে isomorphic হতে বাধ্য। অর্থাৎ $\mathbb{R}^5$-এর ধর্মগুলো জানলেই সেই vector space-টার যাবতীয় ধর্মও জানা হয়ে যাবে। একই রকম কথা যেকোনো field-এর উপর যেকোনো finite dimensional vector space-এর জন্যই বলা যায়। সেখান থেকেই নীচের theorem-টার উদ্ভব--

> **THEOREM**
>
> Any n-dimensional vector space over a field F must be isomorphic to F^n.

তোমরা যারা বড় হয়ে শিক্ষক শিক্ষিকা হবে, তারা এই theorem-টা বিশেষভাবে মনে রেখো। ছাত্রছাত্রীদের প্রথমেই vector space-এর সুদীর্ঘ সংজ্ঞাটা দিয়ে ব্যতিব্যস্ত না করে, খালি F^n-দের নিয়ে কাজ করলেই চলবে। বস্তুতঃ খালি $\mathbb{R}^n$ নিয়ে কাজ করলেই প্রাথমিকভাবে সব finite dimensional vector space-এর রহস্যই বোঝা যায়। সেটা বোঝা হয়ে গেলে অন্যান্য field আর vector space-এর আলোচনা করলেই হয়।

32.1 Isomorphism theorems

এবার আমরা কয়েকটা theorem আলোচনা করব, যাদেরকে একত্রে অনেক সময়ে isomorphism theorems বলা হয়। এদের মূল উদ্দেশ্য হল কিছু isomorphic vector space-দের চট্ করে চিনতে শেখানো। এই theorem-গুলোর মধ্যে প্রথমটাই সবচেয়ে গুরুত্বপূর্ণ। এর মূল ধারণাটা প্রথমে একটা সহজ উদাহরণ দিয়ে বুঝে নিই। কোনো integer-কে যদি 3 দিয়ে ভাগ করো, তবে একটা কিছু remainder (ভাগশেষ) পাবে, যেমন 5-কে ভাগ করলে ভাগশেষ আসবে 2, যদি 9-কে ভাগ করো, তবে আসবে 0 (কারণ ভাগটা মিলে যাবে)। আবার যদি -2-কে ভাগ করো, তবে ভাগশেষ পাবে 1 (কারণ $-2 = 3\times(-1)+1$) ইত্যাদি। অর্থাৎ এই "3 দিয়ে ভাগ করে ভাগশেষ নেওয়া" ব্যাপারটা একটা function দিচ্ছে $f:\mathbb{Z}\to\mathbb{Z}$. যেমন, $f(5)=2$, $f(9)=0$ আর $f(-2)=1$, ইত্যাদি। সংখ্যাটা যাই নাও না কেন, ভাগশেষটা সর্বদাই $0,1,2$-এর মধ্যে কিছু একটা হতে বাধ্য। তাই এই function-টা range হল $\{0,1,2\}$. এবার সেই সব সংখ্যা x ভাবো, যাদের জন্য $f(x)=0$ হয়। এরা হল 3-এর সব multiple-দের set. অর্থাৎ $K=\{0,\pm3,\pm6,...\}$. এবার সেই সব x ভাবো যাদের জন্য $f(x)=1$ হয়। এরা হল $1+K$. একইভাবে যাদের জন্য $f(x)=2$, তারা হল $2+K$. এইভাবে পুরো domain-টা তিনটে গুচ্ছে ভেঙে গেল, যদি প্রতিটা গুচ্ছকে Fig 42-এর মত করে লেবেল লাগানো পোঁটলা হিসেবে কল্পনা করো, তবে তুমি বস্তুতঃ $\{0,1,2\}$ পাচ্ছ, যেটা f-এর image. এই যে ধারণাটা, domain-টাকে গুচ্ছ-গুচ্ছ করে চিন্তা করলে f-এর image-টাই পাওয়া যায়, এইটাই first isomorphism theorem-এর মূল ধারণা। এবার ধারণাটা linear transformation-এর বেলায় কীরকম দাঁড়ায় দেখি।

মনে করো $T:U\to V$ হল একটা linear transformation. ব্যাপারটা দেখিয়েছি Fig 43-এ। ছবিতে $\ker(T)$-কে শেড করে দেখিয়েছি। এবার $\mathrm{Im}(T)$-এর মধ্যে $\vec{0}$ ছাড়া অন্য কিছু নাও, ধরো $\vec{v}$. যেহেতু $\vec{v}\in\mathrm{Im}(T)$, তাই এমন কোনো $\vec{u}\in U$ পাবে যাতে $T(\vec{u})=\vec{v}$ হয়। যেহেতু $\vec{v}\neq\vec{0}$, তাই $\vec{u}\notin\ker(T)$ হবে। এবার দেখো, U-এর যেসব vector-এর উপর T লাগালে $\vec{v}$ হয়, তাদের set-টা হল $\vec{u}+\ker(T)$. কারণ--

> যদি কোনো $\vec{x}\in U$-র জন্য $T(\vec{x})=\vec{v}$ হয়, তবে $T(\vec{x})=T(\vec{u})$ হবে, মানে $T(\vec{x}-\vec{u})=\vec{0}_U$ হবে। তাই $\vec{x}-\vec{u}\in\ker(T)$ হবে। ফলে $\vec{x}=\vec{u}+(\vec{x}-\vec{u})\in\vec{u}+\ker(T)$ হবে।
>
> বিপরীতপক্ষে, যদি $\vec{x}\in\vec{u}+\ker(T)$ হয়, তবে $\vec{x}\in\vec{u}+\vec{w}$, যেখানে $\vec{w}$ হল $\ker(T)$-এর কোনো সদস্য। তাই $T(\vec{x})=T(\vec{u}+\vec{w})=T(\vec{u})+T(\vec{w})=T(\vec{u})+\vec{0}_W=\vec{v}$.

তার মানে ঠিক যেন U-টাকে আমরা পাঁউরুটির মত কতগুলো একইরকম স্লাইসে ভেঙে নিতে পারলাম। তার মধ্যে একটা স্লাইস হল $\ker(T)$, আর বাকিগুলো সেই $\ker(T)$-কেই খানিকটা করে সরিয়ে সরিয়ে তৈরী। Fig 44 দেখে নাও। T লাগালে একেকটা স্লাইস $\mathrm{Im}(T)$-এর একেকটা বিন্দুতে যায়, প্রতি বিন্দুর জন্য ঠিক একটা করে স্লাইস। প্রত্যেকটা স্লাইসকে $\vec{u}+\ker(T)$ আকারে লেখা যায়, যেখানে $\vec{u}\in U$. যদি তুমি যেকোনো দুটো স্লাইস নাও, ধরো $\vec{u}_1+\ker(T)$ আর $\vec{u}_2+\ker(T)$, তবে ওদেরকে যোগ করলে তুমি সর্বদাই $(\vec{u}_1+\vec{u}_2)+\ker(T)$ পাবে। আবার যদি কোনো স্লাইস $\vec{u}+\ker(T)$-কে তুমি কোনো $r\in F$ দিয়ে গুণ করো, তবে তুমি পাবে $r\vec{u}+\ker(T)$ স্লাইসটা। অর্থাৎ যদি U-কে আমরা এক গুচ্ছ স্লাইস বলে ভাবি,

Fig 42

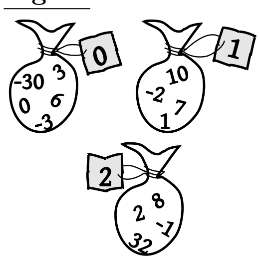


Fig 43

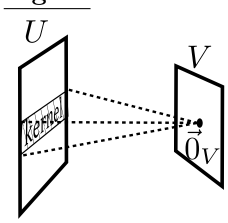
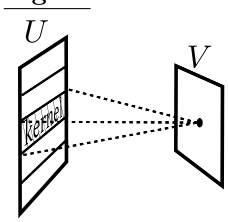


Fig 44



তবে U-টা যেন $\operatorname{Im}(T)$-এর একটা প্রতিরূপ। এই যে বললাম "U-কে একগুচ্ছ স্লাইস বলে ভাবা", অংকের ভাষায় একেই যে $U/\ker(T)$ বলে, সেটা আগেই বলেছি। সেখান থেকে আসছে এই theorem-টা--

> **First isomorphism theorem**
>
> Let $T : U \to V$ be a linear transformation. Then $U/\ker(T) \cong \operatorname{Im}(T)$.

Proof: প্রমাণের কায়দাটা হল সরাসরি একটা isomorphism তৈরী করে দেখানো।

> Shall show that $f : U/\ker(T) \to \operatorname{Im}(T)$ defined as $f(\vec{u} + \ker(T)) = T(\vec{u})$ is an isomorphism.

এটাকে isomorphism দেখানোর পথে চারটে ধাপ--প্রথমে এটা দেখাতে হবে যে, এটা আদৌ একটা ঠিকঠাক function. তার মানেটা বুঝে নাও। আমরা বলেছি $f(\vec{u} + \ker(T)) = T(\vec{u})$. এবার ধরো তুমি এমন একটা $\vec{v} \in U$ পেলে যার জন্য $\vec{u}+\ker(T) = \vec{v}+\ker(T)$ হচ্ছে। তবে তুমি $f(\vec{v}+\ker(T)) = T(\vec{v})$ লিখবে। এদিকে $f(\vec{v}+\ker(T)) = T(\vec{u}+\ker(T))$. কিন্তু সেই একই জিনিসকে আমি বলছি $T(\vec{u})$ আর তুমি বলছ $T(\vec{v})$. এবার যদি দেখা যায় যে, $T(\vec{u}) \neq T(\vec{v})$, তবেই চিত্তির! আমাদের f সেক্ষেত্রে একই জায়গায় দুটো আলাদা value নিয়ে ফেলবে যে! সেটা তো হতে দেওয়া যায় না! তাই f-কে ঠিকঠাক function হবার জন্য সর্বদাই $T(\vec{u}) = T(\vec{v})$ হওয়া দরকার। সেটাই দেখিয়ে নেব প্রথম ধাপে। আমরা ছবিতে যেটাকে বলেছিলাম "একটা গোটা স্লাইস একই বিন্দুতে যাচ্ছে", তারই প্রমাণ এটা--

> Step 1: Well-defined: To show
>
> $\forall \vec{u}, \vec{v} \in U \quad \big(\vec{u} + \ker(T) = \vec{v} + \ker(T) \implies T(\vec{u}) = T(\vec{v})\big).$
>
> $\forall \vec{u}, \vec{v}$ Take any $\vec{u}, \vec{v} \in U$ such that $\vec{u} + \ker(T) = \vec{v} + \ker(T)$.
>
> Then $\vec{u} - \vec{v} \in \ker(T)$.
>
> Hence
>
> $$\begin{aligned} T(\vec{u}) &= T(\vec{v} + (\vec{u} - \vec{v})) \\ &= T(\vec{v}) + T(\vec{u} - \vec{v}) \quad [\because T \text{ is linear}] \\ &= T(\vec{v}) + \vec{0}_V \quad [\because \vec{u} - \vec{v} \in \ker(T).] \\ &= T(\vec{v}), \end{aligned}$$
>
> as required.

এবার দ্বিতীয় ধাপে দেখাব যে, এই f-টা একটা linear transformation.

> Step 2: Linear transformation: Shall show
>
> $\forall x, y \in U/\ker(T) \quad \forall a, b \in F \quad f(ax + by) = af(x) + bf(y).$
>
> $\forall \vec{u}, \vec{v}$ Take any $x, y \in U/\ker(T)$. Then $x = \vec{u}+\ker(T)$ and $y = \vec{v}+\ker(T)$ for some $\vec{u}, \vec{v} \in U$.
>
> $\forall a, b$ Take any $a, b \in F$.

Then

$$\begin{aligned} f(ax+by) &= f((a\vec{u}+b\vec{v})+\ker(T)) \\ &= T(a\vec{u}+b\vec{v}) \quad [\text{ by defn of } f] \\ &= aT(\vec{u})+bT(\vec{v}) \quad [\because T \text{ is linear}] \\ &= af(x)+bf(y). \end{aligned}$$

তৃতীয় ধাপে দেখাব যে, f হল one-one.

Step 3: One-one: Shall show

$\forall x, y \in U/\ker(T) \quad \big(f(x)=f(y) \implies x=y\big).$

$\forall x, y$ Take any $x, y \in U/\ker(T)$ such that $f(x)=f(y)$.

Then $x=\vec{u}+\ker(T)$ and $y=\vec{v}+\ker(T)$ for some $\vec{u}, \vec{v} \in U$.

Since $f(x)=f(y)$, we have $T(\vec{u})=T(\vec{v})$. So $T(\vec{v}-\vec{u})=\vec{0}_V$. Hence $\vec{v}-\vec{u} \in \ker(T)$.

So

$$x=\vec{u}+\ker(T)=\vec{v}+(\vec{u}-\vec{v})+\ker(T)=\vec{v}+\ker(T)=y,$$

as required.

যেহেতু এখানে $\vec{u}-\vec{v} \in \ker(T)$ ছিল, আর $ker(T)$ হল একটা subspace, তাই $\vec{u}-\vec{v}+\ker(T)$ ফের আবার $ker(T)$-ই হল।
শেষ ধাপে দেখাব যে, আমাদের f-টা surjective-ও বটে--

Step 4: Onto: Shall show

$\forall \vec{v} \in \operatorname{Im}(T) \quad \exists x \in U/\ker(T) \quad f(x)=\vec{v}.$

$\forall \vec{v}$ Take any $\vec{v} \in \operatorname{Im}(T)$.

Then $\vec{v}=T(\vec{u})$ for some $\vec{u} \in U$.

$\exists x$ Choose $x=\vec{u}+\ker(T) \in U/\ker(T)$.

Then $f(x)=T(\vec{u})=\vec{v}$, as required.

[Q.E.D]

এই প্রমাণের যুক্তিটা খুব যে কঠিন তা হয়তো নয়, কিন্তু জীবনে প্রথমবার দেখলে হজম হতে সময় নিতে পারে। তাই নীচের অংকটা দিচ্ছি। এখানেও একই যুক্তি আরেকবার লাগবে। এটা যদি করতে পারো, তবে বুঝব যে, যুক্তিটা তোমার সত্যিই হজম হয়েছে।

Exercise 66: Let $T: U \to V$ be a linear transformation. Let W be a subspace of $\ker(T)$. Then prove that $f: U/W \to \operatorname{Im}(T)$ defined as $f(\vec{u}+W)=T(\vec{u})$.is a linear transformation. Is it onto? Is it one-one? ■

Second isomorphism theorem

Let U, V be two subspaces of some vector space. Then

$$U/(U \cap V) \cong (U+V)/V.$$

Exercise 67: এই theorem-টা প্রমাণ করো। তার জন্য এই function-টা নাও $f : U/(U \cap V) \to (U+V)/V$, যেখানে

$$f(\vec{u} + U \cap V) = \vec{u} + V, \quad (\vec{u} \in U).$$

এবার ঠিক আগের theorem-টার প্রমাণের মত করে চার ধাপে এগোও, প্রথমে দেখাও well-defined, তারপর linear transformation, তারপর one-one এবং onto.

HINT:

এখানে একটা জিনিস ধরিয়ে দিই। $U/(U \cap V)$-এর সদস্যদের দেখতে $\vec{u} + U \cap V$-এর মত, যেখানে $\vec{u} \in U$. একইভাবে $(U+V)/V$-এর সদস্যরা হল $\vec{w} + V$-এর মত যেখানে $\vec{w} \in U + V$. এটাকে আরেকটু সহজ করে লেখা যায়। যেহেতু $\vec{w} \in U + V$, তাই $\vec{w} = \vec{u} + \vec{v}$ লেখা যাবে, যেখানে $\vec{u} \in U$ আর $\vec{v} \in V$. যেহেতু V-টা একটা subspace, তাই $\vec{v} + V = V$ হবে। অতএব $(U+V)/V$-এর সদস্যদের দেখতে হবে এইরকম--

$$\vec{w} + V = \vec{u} + \vec{v} + V = \vec{u} + V,$$

যেখানে $\vec{u} \in U$. ■

Third isomorphism theorem

Let V be a vector space. W is a subspace of V, and U is a subspace of W. Then

$$(V/U)/(W/U) \cong V/W.$$

Exercise 68: এই theorem-টা প্রমাণ করো। তার জন্য এই function-টা নাও $f : (V/U)/(W/U) \to V/W$, যেখানে

$$f((\vec{v} + U) + W/U) = \vec{v} + W, \quad (\vec{v} \in V).$$

এবার ফের first isomorphism theorem-টার প্রমাণের মত করে চার ধাপে এগোও, প্রথমে দেখাও well-defined, তারপর linear transformation, তারপর one-one এবং onto.

HINT:

এটা একটু বেশী জটিল দেখতে। যদি বলি $(V/U)/(W/U)$-র সদস্যরা দেখতে কীরকম, বলতে পারবে? এরা সবাই দেখতে $x + W/U$-এর মত, যেখানে $x \in V/U$. সুতরাং x নিজেই আবার $\vec{v} + U$-এর মত দেখতে। সুতরাং সব মিলিয়ে $(V/U)/(W/U)$-এর সদস্যদের চেহারা দাঁড়াচ্ছে $(\vec{v} + U) + W/U$-এর মত, যেখানে $\vec{v} \in V$. ■

Answers

1. প্রথম তিন ক্ষেত্রে হ্যাঁ। শেষের ক্ষেত্রে না।

2. $T(a\vec{x}+b\vec{y}) = T(ax_1+by_1, ax_2+by_2, ax_3+by_3) = \big(2(ax_1+by_1)+(ax_2+by_2)+3(ax_3+by_3),\ 3(ax_1+by_1)-(ax_2+by_2)+(ax_3+by_3),\ -4(ax_1+by_1)+3(ax_2+by_2)+(ax_3+by_3)\big) = a(2x_1+x_2+3x_3,\ 3x_1-x_2+x_3,\ -4x_1+3x_2+x_3)+b(2y_1+y_2+3y_3,\ 3y_1-y_2+y_3,\ -4y_1+3y_2+y_3) = aT(\vec{x})+bT(\vec{y})$.
3. না, হ্যাঁ, না। **4.** (C),(D). **5.** (C). **6.** (D). এখানে আমরা খালি $(x,y,z)\in \operatorname{span}\{(1,1,1),\,(1,1,2),\,(2,2,1)\}$-এর জন্যই $T(x,y,z)$ বার করতে পারি। এই span-এর সকল সদস্যই দেখতে (a,a,b) আকারের। যেহেতু $(1,2,3)$ এই আকারের নয়, তাই $T(1,2,3)$ বার করা যাচ্ছে না।

7. $f(x_1,x_2)=x_1^2$. **8.** Necessary. **9.** $\begin{bmatrix} -3 & 1 & -2 \\ -2 & 3 & 1 \end{bmatrix}$. **10.** $\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ -1 & 0 \end{bmatrix}$.

11. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. **12.** $T(x,y,z)=(-2x-y+6z,\ x-2y+z,\ -x+y+2z)$. এর জন্য matrix-টা হল $\begin{bmatrix} -2 & -1 & 6 \\ 1 & -2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$. **13.** $\begin{bmatrix} 3 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$. **14.** $\begin{bmatrix} 1 & 0 & -\frac{3}{2} \\ 0 & -1 & \frac{3}{2} \\ 0 & 2 & \frac{3}{2} \end{bmatrix}$. **15.** $\begin{bmatrix} 4 & 5 & 5 \\ 2 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix}$.

16. $\begin{bmatrix} 5 & 3 & -2 \\ 0 & 1 & 3 \end{bmatrix}$. **17.** $T(a,b,c)=(a+b+2c, -a+2b+c, b+3c)$. Matrix-টা হল $\begin{bmatrix} 4 & 3 & 2 \\ -2 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix}$. **18.** Explicit representation হল $T(x,y,z)=(x+2y+3z,\ x-2y+2z,\ x-4z)$. আর matrix-টা হল $\begin{bmatrix} -1 & 0.5 & 2.5 \\ -2 & -4.5 & -1.5 \\ 5 & 4.5 & 0.5 \end{bmatrix}$. **19.** $\begin{bmatrix} -9 & -7 & -3 \\ 12 & 9 & 4 \\ 6 & 4 & 3 \end{bmatrix}$. **20.** $\frac{1}{2}\begin{bmatrix} -1 & 4 & 3 \\ 3 & 4 & -1 \\ 3 & -4 & 7 \end{bmatrix}$.

21. $\begin{bmatrix} 4 & 3 & 3 \\ -4 & -2 & -1 \end{bmatrix}$. **22.** Explicit representation হল $T(x,y,z)=\big(2x+2y+z, \frac{x-y+3z}{2}\big)$. আর matrix-টা হল $\begin{bmatrix} -1 & -4 & -2 \\ 4 & 8 & 5 \end{bmatrix}$. **23.** $\begin{bmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}$. আগের matrix-টার inverse এটা।

24. $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 4 & 7 \end{bmatrix}$. **25.** $B_1^{-1}B_2$. **26.** $\begin{bmatrix} -1.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix}$. হ্যাঁ, ফেরা যাচ্ছে। **27.** $\mathcal{D}=\{2-x,1\}$.
28. $P_{ED}^{-1}MP_{BA}$. **29.** $\operatorname{Im}(T)=\operatorname{span}\{(1,0,1),(0,1,-1)\}$. এর প্রতিটা vector-ই হল $(a,b,a-b)$ জাতীয়। তাই $(1,2,3)\notin \operatorname{Im}(T)$, কিন্তু $(1,2,-1)\in\operatorname{Im}(T)$. **30.** 3. **31.** (A). এর ব্যাখ্যাটা এইরকম--

$$\begin{aligned}&(x,y,z)\in\ker(T) \iff (x+y, y-z, z+x)=(0,0,0)\\ &\iff x=-y=-z \iff (x,y,z)\in\operatorname{span}\{(1,-1,-1)\}.\end{aligned}$$

32. $\{\vec{0}\}$. **33.** $T(x,y,z)=(y+z,x+z,x+y)$. এখানে $\ker(T)=\{(0,0,0)\}$ এবং $\operatorname{Im}(T)=\mathbb{R}^3$.
34. (B),(D). **35.** $(1,0,0)\in\mathscr{C}(A)$, কিন্তু $(1,0,0)\notin\mathscr{C}(B)$. **36.** না। যেমন ধরো $T:\mathbb{R}^2=\mathbb{R}$ নিলাম $T(x,y)=(x,0)$. তবে $\{\vec{e}_1,\vec{e}_2\}$ এবং $\{\vec{e}_1,\vec{e}_2\}$-এর সাপেক্ষে matrix-টা হবে $A=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. কিন্তু যদি domain-এর basis-টা $\{\vec{e}_2,\vec{e}_1\}$ নিতাম matrix-টা হয়ে যেত $B=\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. এখানে $\mathscr{N}(A)\neq\mathscr{N}(B)$, কারণ $\begin{bmatrix} 0 \\ 1 \end{bmatrix}\in\mathscr{N}(A)$, কিন্তু $\begin{bmatrix} 0 \\ 1 \end{bmatrix}\notin\mathscr{N}(B)$. **38.** 100. এখানে image-টা সেই সব polynomial-এর নিয়ে তৈরী, যাদের degree≤ 99. **39.** $T(x,y,z)=(y+z,x+z,x+y)$. এখানে $\ker(T)=\{\vec{0}\}$ এবং $\operatorname{Im}(T)=\mathbb{R}^3$. তাই $\dim(\ker(T))+\dim(\operatorname{Im}(T))=0+3=3$. **40.**

Let $(x_1 x_2, x_3, x_4) \in \ker(T)$.
Then

$$\begin{aligned} 3x_1 - 2x_2 &= 0, \\ -x_3 - x_4 &= 0, \\ x_1 + x_2 - 2x_3 - 3x_4 &= 0. \end{aligned}$$

Solving we have

$$x_1 = \text{arbitrary},\ x_2 = \tfrac{3}{2}x_1, \quad x_3 = -\tfrac{5}{2}x_1, \quad x_4 = \tfrac{5}{2}x_1.$$

Thus a vector is in $\ker(T)$ if and only if it is of the form

$$\begin{bmatrix} x_1 \\ \frac{3}{2}x_1 \\ -\frac{5}{2}x_1 \\ \frac{5}{2}x_1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ \frac{3}{2} \\ -\frac{5}{2} \\ \frac{5}{2} \end{bmatrix}$$

for some $x_1 \in \mathbb{R}$.
Thus a basis of $\ker(T)$ is $\{(1, \frac{5}{2}, -\frac{5}{2}, \frac{5}{2})\}$ and the nullity is 1.
So $\text{rank}(T) = \dim(\mathbb{R}^4) - \text{nullity}(T) = 4 - 1 = 3$.

41.

Here

$$\begin{aligned} & (x_1, x_2, x_3) \in \ker(T) \\ \iff & (x_1 + x_2 + x_3,\ 2x_1 + x_2 + 2x_3,\ x_1 + 2x_2 + x_3) = (0, 0, 0), \end{aligned}$$

ie,

$$\begin{aligned} x_1 + x_2 + x_3 &= 0, \\ 2x_1 + x_2 + 2x_3 &= 0, \\ x_1 + 2x_2 + x_3 &= 0. \end{aligned}$$

Solving,

$$x_1 = \text{ arbitrary},\ x_2 = 0, \quad x_3 = -x_1.$$

Thus

$$\ker(T) = \{x_1(1, 0, -1) \ : \ x_1 \in \mathbb{R}\}.$$

$\therefore \{(1, 0, -1)\}$ is a basis of $\ker(T)$.
We know that

$$\dim(\text{Im}(T)) + \dim(\ker(T)) = \dim(\mathit{domain}(T)).$$

> So $\dim(\mathrm{Im}(T)) = 3 - 1 = 2$.

42. $f(x) = (x, 0)$. **43.** অসম্ভব। **44.** অসম্ভব। **45.** $f(x, y) \equiv 0$. **46.**

> Let $T(x, y, z) = (p, q, r)$. Then
>
> $$\begin{aligned} 3x + z &= p, \\ -2x + y &= q, \\ -x + 2y + 4z &= r. \end{aligned}$$
>
> Solving for x, y, z in terms of p, q, r, we get
>
> $$x = -\frac{r - 2q - 4p}{9}, y = -\frac{2r - 13q - 8p}{9}, z = \frac{r - 2q - p}{3}.$$
>
> Thus, for each $(p, q, r) \in \mathbb{R}^3$ there exists a unique $(x, y, z) \in \mathbb{R}^3$ such that $T(x, y, z) = (p, q, r)$.
>
> Hence $T : \mathbb{R}^3 \to \mathbb{R}^3$ is invertible.

47.

> Let $\dim(V) = n$ and $\dim(W) = m$. Since $A_{m\times n}$ is full row rank, hence $r(A) = m$.
>
> So $\dim(\mathrm{Im}(T)) = m$. But $\mathrm{Im}(T) \subseteq W$. Since $\dim(W) = m$, hence $\mathrm{Im}(T) = W$, as required.

48. $\vec{v}_1 = (1, 0)$ আর $\vec{v}_2 = (0, 1)$ নেওয়া যেতে পারে। **49.** $\{(1, 0), (0, 1)\}$. **50.**

> For $(x_1, x_2, x_3) \in \mathbb{R}^3$
>
> $$\begin{aligned} & (x_1, x_2, x_3) \in \ker(T) \\ \iff \quad & T(x_1, x_2, x_3) = (0, 0, 0, 0) \\ \iff \quad & (x_2 + x_3,\ \ x_3 + x_1,\ \ x_1 + x_2, x_1 + x_2 + x_3) = (0, 0, 0, 0), \end{aligned}$$
>
> ie, if and only if $x_1 = x_2 = x_3 = 0$.
>
> So $\ker(T) = \{(0, 0, 0)\}$.
>
> So T must be injective.

আমরা জানি যে, কোনো linearly independent set-এর উপর একটা injective linear transformation লাগালে ফের linearly independent set-ই পাওয়া যায়। তাই--

> Hence we conclude that the image of the given set must be linearly independent.

ছোটো করে প্রমাণটা ঢুকিয়েই দিই--

⟦ Because:

The given set $\{e_1, e_2, e_3\}$ is linearly independent.
If $aT(e_1) + bT(e_2) + cT(e_3) = (0, 0, 0, 0)$, then
$T(ae_1 + be_2 + ce_3) = (0, 0, 0, 0)$,
ie, $ae_1 + be_2 + ce_3 \in \ker(T) = \{(0, 0, 0)\}$.
So, by independence of $\{e_1, e_2, e_3\}$, we have $a = b = c = 0$.

⟧

51. $\dim(T(U)) \leq \dim(U)$ দেখানোর জন্য ধরো $\{\vec{v}_1, ..., \vec{v}_k\}$ হল $T(U)$-এর একটা basis. তার মানে $\dim(T(U)) = k$ ধরেছি। যেহেতু $\vec{v}_i \in T(U)$, তাই এমন $\vec{u}_i \in U$ পাবে যার জন্য $\vec{v}_i = T(\vec{u}_i)$ হবে। তবে $\{\vec{u}_1, ..., \vec{u}_k\}$ অবশ্যই linearly independent হবে। যেহেতু U-এর মধ্যে k সাইজের একটা linearly independent আছে, তাই $\dim(U) \geq k$ হতে বাধ্য, কারণ dimension মানে হল সবচেয়ে বড় linearly independent set-এর সাইজ। **52.** হ্যাঁ। **53.** ভুল। V-এর basis-এর image-টা $\text{Im}(T)$-এর একটা basis হবে। কিন্তু T যদি surjective না হয়, তবে $\text{Im}(T)$ হবে W-এর চেয়ে ছোটো। তাই $\text{Im}(T)$-এর basis-টা কোনোভাবেই সেক্ষেত্রে পুরো W-এর basis হতে পারবে না। Counterexample-এর জন্য এমন কোনো linear transformation নিতে হবে, যেটা injective হলেও surjective নয়। যেমন--

Let $F = \mathbb{R}, V = \mathbb{R}^2, \quad W = \mathbb{R}^3$ and

$$T(x, y) = (x, y, 0).$$

Here $T : V \to W$ is an injective linear transformation.

Now $\{(1, 0), (0, 1)\}$ is a basis of V, but

$$\{T(1, 0), T(0, 1)\} = \{(1, 0, 0), (0, 1, 0)\}$$

is not a basis in W.

54. হ্যাঁ, কারণ $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ যেহেতু linearly independent, তাই $\{\vec{v}_1, ..., \vec{v}_n\}$-ও তাই হবে। আবার T যেহেতু bijective, তাই surjective-ও বটে। ফলে $\text{Im}(T) = W$. সুতরাং $\{T(\vec{v}_1), ..., T(\vec{v}_n)\}$ হল $\text{Im}(T)$-এর একটা spanning set. যেহেতু T-টা injective-ও বটে, তাই $\{\vec{v}_1, ..., \vec{v}_n\}$ অবশ্যই V-এর spanning set হবে।
55. এই অংকটা matrix দিয়ে সুন্দর করে করা যায়।

Let $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ be a basis of $\mathbb{R}^3$. Then $B = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$ is a nonsingular matrix.

$$\text{Now } T(x_1, x_2, x_3) = \underbrace{\begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 4 \\ 1 & -1 & 3 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

If $T(x_1, x_2, x_3) = (0, 0, 0)$, then by solving the equations we have $(x_1, x_2, x_3) = (0, 0, 0)$.

এখানে equation তিনটের সমাধান দেখালাম না, তুমি ঢুকিয়ে দিও।

> So $\text{nullity}(A) = 0$ and hence A is nonsingular, since A is square.
>
> The image of the basis consists of the columns of AB.
>
> Since A and B are both nonsingular, AB is also nonsingular. So the columns form a basis, as required.

56. Abelian group তো আগেই দেখিয়েছি। Function composition যে associative সেটা একটা standard result. পড়ে থাকে খালি distributive দেখানো। ধরো $f, g, h \in L(V, V)$. দেখাব যে $f \circ (g + h) = f \circ g + f \circ h$. এটা হবে কারণ $\forall \vec{v} \in V \quad (f \circ (g + h))(\vec{v}) = f(g(\vec{v}) + h(\vec{v})) = f(g(\vec{v})) + f(h(\vec{v}))$, কারণ f হল linear. একইভাবে $(g + h) \circ f = g \circ g + h \circ f$-ও প্রমাণ হবে। **57.** (A). **58.** (B). **59.** ধরো $f : V \to W$ আর $g : W \to V$ দুজনেই injective linear transformation. তাহলে $\text{nullity}(f) + \dim(\text{Im}(f)) = \dim(V)$ আর $\text{nullity}(g) + \dim(\text{Im}(g)) = \dim(W)$. যেহেতু $\text{nullity}(f) = \text{nullity}(g) = 0$, তাই $\dim(\text{Im}(f)) = \dim(V)$ আর $\dim(\text{Im}(g)) = \dim(W)$. কিন্তু $\text{Im}(f)$ হল W-এর subspace, আর $\text{Im}(g)$ হল V-এর subspace. সেই কারণে $\dim(\text{Im}(f)) \leq \dim(W)$ আর $\dim(\text{Im}(g)) \leq \dim(V)$. তাই $\dim(W) \leq \dim(V)$ এবং একইসঙ্গে $\dim(V) \leq \dim(W)$. অতএব $\dim(V) = \dim(W)$. **60.** যদি V, W দুজনেই একই field-এর উপরে দুটো vector space হয়, যাদের dimension সমান (এবং finite), আর $f : V \to W$ হয় যেকোনো linear transformation, তবে f-টা injective হলে surjective-ও হতে বাধ্য। প্রমাণটা এইরকম-- $\text{rank}(f) + \text{nullity}(f) = \dim(V) = \dim(W)$. এদিকে injective বলে $\text{nullity}(f) = 0$. সুতরাং $\text{rank}(f) = \dim(W)$. তার মানে $\dim(\text{Im}(f)) = \dim(W)$. যেহেতু $\text{Im}(f)$ হল W-এর subspace, তাই $\text{Im}(f) = W$ হয়ে যাচ্ছে, মানে f-টা surjective হচ্ছে।
61. Surjective তো দিয়েই দিয়েছে, খালি injective দেখালেই হবে, মানে $\text{nullity}(T) = 0$ দেখাব। এখন $\text{nullity}(T) = \dim(W) - \dim(\text{Im}(f)) = \dim(W) - \dim(W) = 0$, যেহেতু $\text{Im}(f) = W$.
62. $\text{nullity}(f) + \dim(\text{Im}(f)) = \dim(V)$ আর $\text{nullity}(f) = 0$. এদিকে $\text{Im}(f)$ হল W-এর একটা subspace, তাই $\dim(\text{Im}(f)) \leq \dim(W)$. সুতরাং $\dim(V) \leq \dim(W)$. আবার $\text{nullity}(g) + \dim(\text{Im}(g)) = \dim(V)$. আর $\text{Im}(g) = W$. সুতরাং $\text{nullity}(g) + \dim(W) = \dim(V)$. যেহেতু $\text{nullity}(g) \geq 0$, তাই $\dim(W) \leq \dim(V)$. সুতরাং পেলাম $\dim(V) = \dim(W)$. এ থেকে আরও পেয়ে গেলাম যে $\dim(\text{Im}(f)) = \dim(W)$, তাই $\text{Im}(f) = W$, মানে f-টা surjective, আর তাই invertible (কারণ injective তো দেওয়াই ছিল!)। একইভাবে $\text{nullity}(g) = 0$, তাই g-টা injective, অতএব invertible-ও। **63.** $T(x, y, z) = (a, b, c)$ নিয়ে শুরু করলে তিনটে equation পাবে $x + y = a,\ y + z = b$ এবং $x + z = c$. এদেরকে solve করলে আসবে $x = (a + b + c)/2 - b$, $y = (a + b + c)/2 - c$ এবং $z = (a + b + c)/2 - a$. যেহেতু প্রতিটা $(a, b, c) \in \mathbb{R}^3$-এর জন্যই ঠিক একটাই (x, y, z) পাওয়া গেল, তাই $T(x, y, z)$-টা one-one এবং onto হতে বাধ্য।
এই সমাধানে আমরা এটা ব্যবহার করলাম না যে, $T(x, y, z)$ একটা linear transformation. সেটা ব্যবহার করলে একটু সহজে হত। $T(x, y, z) = (0, 0, 0)$ মানে $(x, y, z) = (0, 0, 0)$. তার মানে $\ker(T) = \{(0, 0, 0)\}$. সুতরাং one-one. আবার $\dim(\text{Im}(T)) = 3 - \dim(\ker(T)) = 3 - 0 = 3$. সুতরাং onto.
64. $T^{-1}(x, y, z) = (x/2, (x - y)/5, z - 2(x - y)/5)$. **65.**

> Let $T(\vec{x}) = A\vec{x}$. Then
>
> $$A \underbrace{\begin{bmatrix} 1 & 3 & 0 \\ 2 & 0 & 3 \\ 3 & 1 & 1 \end{bmatrix}}_{P} = \underbrace{\begin{bmatrix} -3 & -5 & 4 \\ 0 & 2 & -1 \\ -2 & -2 & 1 \end{bmatrix}}_{Q}.$$
>
> Now P is nonsingular.

⟦ Because:

The first two rows are not multiples of each other.
Let, if possible, $(3,1,1) = a(1,3,0) + b(2,0,3)$ for some $a, b \in \mathbb{R}$.
Then $a = b = \frac{1}{3}$. But $3 \neq a + 2b$. So $(3,1,1) \notin \operatorname{span}\{(1,3,0), (2,0,3)\}$.

⟧

Also Q is nonsingular.

⟦ Because:

$(-5,2,-2), (4,-1,1)$ are not multiples of each other.
Also any vector in their span must be of the form $(a, b, -b)$ for some $a, b \in \mathbb{R}$. But $(-3,0,-2)$ is not of this form.

⟧

Hence $A = QP^{-1}$. Since product of nonsingular matrices is again nonsingular, A is nonsingular.

So T is an isomorphism, as required.

66. $f\big(a(\vec{u}+W) + b(\vec{v}+W)\big) = f\big(a\vec{u} + b\vec{v}) + W\big) = T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v}) = af(\vec{u}+W) + bf(\vec{v}+W)$. তাই linear transformation হবে। Onto, কিন্তু one-one নাও হতে পারে।

Chapter V
Determinant

Linear algebra-র কোনো কোনো নতুন সিলেবাসে determinant-এর উল্লেখ নেই। কিন্তু সিলেবাসকর্তারা যাই বলুন, determinant জিনিসটা linear algebra-র একটা গুরুত্বপূর্ণ ধারণা। আর বিভিন্ন MCQ-কে ঘায়েল করতেও determinant প্রায়শঃই কাজে লাগে। সেই সব কারণেই এই অধ্যায়টা বইতে রেখেছি। সিলেবাসের বাইরে যেতে যাদের বড়ই আপাত্তি, তাদেরও বলব এই অধ্যায়ের প্রথম পাঁচ দিনের পড়াটায় অন্ততঃ একবার হালকাভাবে চোখ বুলিও।

DAY 33
গোড়ার কথা

এই অধ্যায়ে আমরা এমন একটা জিনিস শিখব, যার ইতিহাস অতি বিচিত্র। জিনিসটা হল **determinant**, যার সঙ্গে কিছু পরিচয় হায়ার সেকণ্ডারীতেই হয়ে থাকে। সংক্ষেপে বললে--কোনো square matrix দেওয়া থাকলে তা থেকে বিশেষ কায়দায় একটা সংখ্যা বার করা যায়, সেই সংখ্যাটাকে বলে matrix-টার determinant. এই বর্ণনাটুকুর মধ্যে অদ্ভুত কিছু নেই, সমস্যা হল ওই "বিশেষ কায়দা"-টা নিয়ে। সেটা দেখতে এতই অদ্ভুত যে, প্রথমে দেখলে থতমত খেয়ে ভাবতে হয় "এটা আবার কোথা থেকে এল রে বাবা?" যেমন ধরো, যদি matrix-টা 2×2 হয়--

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

তবে তার determinant হবে $|A| = ad - bc$. কঠিন কিছু নয়, কিন্তু হঠাৎ খামোখা কেন লোকে এই জিনিসটা নিয়ে মাথা ঘামাবে, তার সদুত্তর চট্ করে পাওয়া যায় না। যদি matrix-টা 3×3 হত,

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix},$$

তবে determinant-এর ফর্মুলাটা হত

$$|A| = aei - bdi - afh + cdh + bfg - ceg.$$

এটাও বিরাট কিছু কঠিন জিনিস নয়, খালি কয়েকটা সংখ্যা গুণ আর যোগ করে তৈরী। কিন্তু বেশ লম্বা জিনিস, তাই বার করতে খুব মজা লাগবে এমন নয়। যদি 4×4 সাইজের matrix নিয়ে কাজ করতাম, তবে ফর্মুলাটা আরও লম্বা হত, বার করা তো দূরের কথা, ফর্মুলাটা লিখতেই হাত টনটন করত। যদি আরো বড় square matrix নিতে, ধরো 10×10, তবে খালি ফর্মুলাটা লিখতেই কয়েক পাতা লাগত। কথা হল এই ফর্মুলাগুলো এল কী করে? এদের দিয়ে হয়টাই বা কী?

এর একটা ইতিহাস আছে। কিন্তু সেই প্রসঙ্গে যাওয়ার আগে বলে রাখি যে, determinant মোটামুটিভাবে একটা কাজেই লাগে, এবং সেটা হল--যদি A একটা square matrix হয়, তবে তার nonsingular হওয়া আর determinant $\neq 0$ হওয়া একই ব্যাপার। এই কথাটা আমরা যথাসময়ে প্রমাণ করব, আপাততঃ ছোট্ট করে ইতিহাসের গল্পটা সেরে নিই।

Leibnitz (লাইবনিৎস) নামে একজন জার্মান গণিতজ্ঞের কথা তোমরা হয়তো শুনেছ, যিনি ক্যালকুলাসের একজন আবিষ্কর্তা। তিনি কোনো একসময়ে এই তিনটে equation-এর সমাধান করছিলেন--

$$\begin{aligned} 10+11x+12y &= 0, \\ 20+21x+22y &= 0, \\ 30+31x+32y &= 0. \end{aligned}$$

এখানে দুটো unknown আছে, অথচ equation আছে তিনটে। সমাধানটা কত হবে সেটা বড় প্রশ্ন ছিল না, প্রশ্ন ছিল আদৌ সমাধান আছে কিনা। এই কথাটা যদি স্পষ্ট না হয়, তবে প্রথমে এই সহজ তিনটে equation দ্যাখো--

$$\begin{aligned} x &= 1, \\ y &= 2, \\ x+y &= 5. \end{aligned}$$

বুঝতেই পারছ এমন কোনো x, y পাওয়া অসম্ভব, যারা একই সঙ্গে এই তিনটে equation-কেই satisfy করে, কারণ যদি $x = 1$ আর $y = 2$ হয়, তবে $x + y$ হবে 3, সেটা কোনোভাবেই 5 হতে পারে না। Leibnitz-ও একইভাবে দেখতে চাইছিলেন--এমন কোনো x, y সম্ভব কিনা, যাতে তাঁর তিনটে equation-ই satisfied হয়। এটা করার সহজ কায়দা হল যেকোনো দুটো equation নিয়ে তা থেকে x, y বার করা, তারপর সেই value-দুটো তৃতীয় equation-এ বসিয়ে দেখা মিলছে কিনা। কিন্তু Leibnitz চাইছিলেন বেশ একটা চট্‌জলদি কায়দা, যেটা দিয়ে কোনো কিছু সমাধান না করেই বলে দেওয়া যাবে সমাধান আছে কি নেই। এর জন্য তিনি প্রথমে সংখ্যার বদলে $a, b, c,$ ইত্যাদি লিখে নিলেন--

$$ax + by + c = 0 \tag{1}$$
$$dx + ey + f = 0 \tag{2}$$
$$gx + hy + i = 0. \tag{3}$$

পরিকল্পনাটা হল এটাকে in general একবার solve করে দেখা যে, কোথায় কী গোলমাল হলে সমাধান থাকবে না। তাহলে প্রতিবার আর আলাদা করে solve করতে হবে না, খালি সেই গোলমালটুকু আছে কিনা সেটা গোড়াতে পরীক্ষা করে নিলেই চলবে। তিনি প্রথম দুটো equation থেকে x-কে কাটাকাটি করে সরিয়ে দেওয়ার জন্য $d \times (1) - a \times (2)$ করলেন--

$$(bd - ae)y + (cd - af) = 0.$$

এইখান থেকে ফস্ করে $y = -\frac{cd-af}{bd-ae}$ লিখে দেওয়ার লোভ হয় বটে, কিন্তু যেহেতু আমরা জানি না $bd - ae \neq 0$ কিনা, তাই সে লোভটা সংবরণ করাই ভালো। এইবার Leibnitz প্রথম আর তৃতীয় equation-টা থেকে x-কে সরিয়ে দিলেন একইভাবে, মানে $g \times (1) - a \times (3)$ করলেন। তাতে পেলেন--

$$(bg - ah)y + (cg - ai) = 0.$$

এবার প্রশ্ন হল-- এই যে দুটো equation পেলাম y-এর জন্য, এরা কি একই সঙ্গে ঠিক হতে পারে? পারে, যদি equation-দুটো একটা আরেকটার multiple (গুণিতক) হয়, যেমন একটা হয়তো $2y + 3 = 0$ আর অন্যটা $4y + 6 = 0$. এইরকম multiple হওয়ার শর্তটা লেখার একটা কায়দা হতে পারে এইরকম--

$$\frac{bd - ae}{bg - ah} = \frac{cd - af}{cg - ai},$$

এখানে বাঁদিকটা হল y-এর coefficient দুটোর ratio, আর ডানদিকটা constant term দুটোর ratio. কিন্তু তাতে আবার denominator-দুটো শূন্য হয়ে যাওয়া নিয়ে দুশ্চিন্তা করতে হয়। তাই Leibnitz ভাগের পথে না গিয়ে শর্তটাকে গুণ দিয়ে লিখলেন, এইভাবে--

$$(bd - ae)(cg - ai) = (cd - af)(bg - ah).$$

দেখতেই পাচ্ছ যে Leibnitz বেশ সাবধানে এগিয়েছিলেন, যাতে কোথাও কিছু একটা শূন্য হয়ে গেলেও ঝামেলা না হয়। খালি একটা জায়গাতে কিন্তু সেটা তিনি আটকাতে পারেন নি। সেটা হল একেবারে প্রথম ধাপে। তিনি দুজায়গায় x-কে তাড়িয়েছেন, দুবারই প্রথম equation-টা ব্যবহার করে। সুতরাং এইখানে ধরে নিয়েছেন যে, প্রথম equation-টাতে x আছে, অর্থাৎ $a \neq 0$. অবশ্যই Leibnitz এই কথাটা মাথায় রেখেছিলেন। এইবার সেই কথাটা কাজে দেবে। তিনি শর্তটাকে বড় করে লিখলেন, কিছু কাটাকাটি যায় কিনা দেখার জন্য--

$$bcdg - abdi - acde + a^2ei = bcdg - acdh - abfg + a^2fh.$$

দুদিক থেকেই $bcdg$ কাটা যায়, এবং বাকিগুলো থেকে a-কে কমন নেওয়া যায়--

$$a(-bdi - cde + aei) = a(-cdh - bfg + afh).$$

আগেই বলেছি যে, এখানে $a \neq 0$ ধরে কাজ হচ্ছে, সুতরাং দুদিক থেকে a-কে কেটে নেওয়া যায়, তাহলে যে শর্তটা পাচ্ছি সেটা হল $-bdi - cde + aei = -cdh - bfg + afh$, বা

$$-bdi - cde + aei + cdh + bfg - afh = 0. \qquad (*)$$

তুমি হয়তো এটার দিকে এক ঝলক তাকিয়েই বলে উঠবে যে, বাঁদিকের জিনিসটাই তো determinant! কিন্তু Leibnitz-এর পক্ষে সেটা বলা সম্ভব ছিল না, কারণ তিনি তো আর determinant-এর নাম শোনেন নি। বস্তুতঃ এই শর্তটার চেহারা দেখে পুলকিত হবার মত কোনো কারণই তাঁর কাছে ছিল না। তাঁর চিন্তা ছিল অন্য জায়গায়--তিনি চাইছিলেন equation-গুলোকে in general সমাধান করতে, কিন্তু $a \neq 0$ ধরতে হওয়ায় ব্যাপারটা পুরো general থাকছে না। সুতরাং যদি $a = 0$ হয়, তবে কী করা উচিত তিনি এবার তাই নিয়ে পড়লেন। সেক্ষেত্রে অন্য কোনো একটা equation ব্যবহার করে বা অন্য কোনো একটা variable-কে তাড়িয়ে শুরু করতে হবে। যেমন ধরো যদি $e \neq 0$ হয়, তবে দ্বিতীয় equation-এ y আছে, সেটাকে ব্যবহার করে সব y-কে তাড়ানোর চেষ্টা করলে ব্যাপারটা দাঁড়াবে এইরকম--প্রথমে $e \times (1) - b \times (2)$ করে

$$(ae - bd)x + (ce - bf) = 0$$

পাব, এবং $h \times (2) - e \times (3)$ করে পাব

$$(dh - eg)x + (fh - ei) = 0.$$

এই দুটো equation একে অপরের multiple হবে যদি

$$(fh - ei)(ae - bd) = (ce - bf)(dh - eg)$$

হয়। কাটাকাটি করে দেখো, সেই (*)-এর শর্তটাই আবার পাবে! এইভাবে তুমি যে equation থেকে যাই variable সরাও না কেন, সব ক্ষেত্রেই ঘুরে ফিরে সেই একই শর্ত এসে উপস্থিত হয়! বিশ্বাস না হয় নিজেই করে দেখো (যেমন $c \neq 0$ ধরে প্রথম equation-টা ব্যবহার করে z-কে তাড়াও)!

যেদিকে থেকেই এগোই, সর্বদাই সেই একই শর্ত! এই কাণ্ড দেখে Leibnitz তাজ্জব হয়ে ভাবলেন যে, শর্তটার কিছু একটা ব্যাপার আছে, নইলে যেদিক দিয়েই শুরু করি একই জিনিস আসছে কেন! তিনি ভারী ব্যস্ত হয়ে তাঁর এই সদ্য-আবিষ্কৃত কায়দার কথা লিখে পাঠালেন L'Hôpital-কে (উচ্চারণ হল লোপিতাল। অনেকে ভুল করে এল হস্পিটাল, বা এল হোপিটাল বলে ফেলে)। এই L'Hôpital ভদ্রলোক কোনো গণিতজ্ঞ নন, ইনি ছিলেন ফ্রান্সের একজন জমিদার, যিনি অংক ভালোবাসতেন, এবং গণিতজ্ঞদের পৃষ্ঠপোষকতা করতেন। বস্তুতঃ L'Hôpital-কে এই চিঠিটা Leibnitz লিখেছিলেন বলেই আমরা এত কথা জানতে পেরেছি।

যাই হোক, সেই হল determinant-এর শুরু। সেটা 1683 খ্রীষ্টাব্দ। আশ্চর্যের কথা হল, ঠিক সেই একই সালে আরেক জন গণিতজ্ঞের লেখাতেও determinant-এর দেখা পাওয়া যায়। ইনি একজন জাপানী, নাম সেকি। বলাই বাহুল্য যে, সেই প্রাচীন কালে Leibnitz এবং সেকির পক্ষে পরস্পরের কাজের খবর জানা সম্ভব ছিল না। তখনো অবশ্য determinant নামটা হয় নি, Leibnitz একে বলতেন resultant.

3×3 matrix-এর ক্ষেত্রে শর্তটা বার করার পর স্বাভাবিকভাবেই জানতে ইচ্ছা হয় 4×4 বা 5×5 হলে সেরকম কিছু বেরোয় কিনা। Leibnitz সে চেষ্টা করে বুঝলেন যে কাজটা করা যাচ্ছে, কিন্তু $a, b, c, d,$ ইত্যাদি ধরে সমাধান করার কায়দাটা

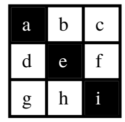


Fig 1

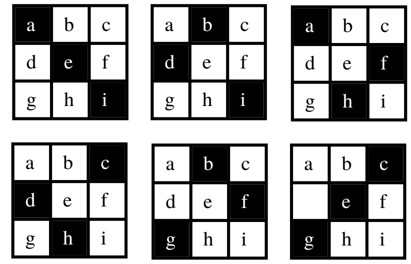


Fig 2

বড্ড কঠিন হয়ে পড়ছে। সুতরাং পুরো জিনিসটাকে একবারে গুছিয়ে লিখে দেওয়া যায় কিনা তার চেষ্টা করতে লাগলেন। এই চেষ্টা করতে গিয়ে কী পেলেন সেটা এবার বলি।
প্রথমে 3×3 উদাহরণটা নাও। আমাদের matrix-টা ছিল--

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}.$$

তাহলে এর determinant পেয়েছিলাম

$$aei + cdh + bfg - afh - bdi - cde.$$

প্রথমেই দ্যাখো determinant-টা অনেকগুলো জিনিস যোগ-বিয়োগ করে তৈরী। প্রত্যেকটা জিনিস তৈরী হয়েছে matrix-টার বিভিন্ন entry গুণ করে করে। যদি 2×2 matrix নিতাম $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, তবে determinant হত $ad-bc$. মানে প্রত্যেকটা term তৈরী হচ্ছে দুটো করে entry গুণ করে। 3×3-এর বেলায় তিনটে করে। সুতরাং আন্দাজ করা যাচ্ছে যে $n\times n$-এর ক্ষেত্রে n-খানা করে গুণ হবে। এরপর দেখতে হবে এরকম কতগুলো জিনিস যোগ-বিয়োগ হচ্ছে। 2×2-এর বেলায় দুটো, 3×3-এর বেলায় 6-টা। 4×4-এর বেলায় Leibnitz দেখলেন 24-টা জিনিস যোগ-বিয়োগ করতে হচ্ছে। সুতরাং আন্দাজ করতে পারলেন যে $n\times n$-এর বেলায় $n!$-খানা জিনিস থাকবে। এইবার তিনি খুঁটিয়ে দেখতে বসলেন যে, কোন্ কোন্ entry-গুলো গুণ হচ্ছে। 3×3-এর উদাহরণটা দেখা যাক। প্রথমেই আছে aei. এই তিনটেকে কালো করে দিলাম Fig 1-এ। এইভাবে সবগুলো থেকে পেলাম Fig 2. এই ছবিটা থেকে বেশ একটা প্যাটার্ণ পাওয়া যাচ্ছে। প্রথম কথা হল কখনোই একই row বা column থেকে একাধিক entry নেওয়া হচ্ছে না। প্রত্যেকটা গুণফলে প্রত্যেকটা row থেকে ঠিক একটা করেই entry আছে, এবং প্রত্যেকটা column থেকেও তাই। এইটা দেখে Leibnitz-এর মাথায় একটা বুদ্ধি খেলে গেল। তিনি বুঝলেন যে $a, b, c, d,$ ইত্যাদি না লিখে row, column-এর সংখ্যা দিয়ে লিখলে বেশী সুবিধা হবে। ঠিক কিভাবে লিখলে সবচেয়ে সুবিধা হবে, তা নিয়ে তিনি বিস্তর গবেষণা করেছিলেন। বস্তুতঃ তাঁর অপ্রকাশিত রচনাবলীর মধ্যে পঞ্চাশেরও বেশী লেখার পদ্ধতির হদিশ পাওয়া যায়!! যাই হোক আমরা আধুনিক notation অনুযায়ী $a_{11}, a_{12},$ ইত্যাদি লিখব--

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

তাহলে প্রতিটা গুণফলই দেখতে হবে এইরকম--

$$a_{1i}a_{2j}a_{3k},$$

যেখানে i, j, k-রা সবাই আলাদা আলাদা হবে, কারণ কোনো column থেকে একাধিক entry আসতে পারে না। এদিকে i, j, k-র মধ্যে $1, 2, 3$-এর প্রত্যেককেই আসতে হবে। সুতরাং কতভাবে i, j, k নেওয়া যায়? উত্তর হল $1, 2, 3$-কে যতভাবে সাজিয়ে লেখা যায়, মানে $3! = 6$ ভাবে--

$$\begin{matrix} 1 & 2 & 3, & \qquad & 1 & 3 & 2, \\ 2 & 1 & 3, & & 2 & 3 & 1, \\ 3 & 1 & 2, & & 3 & 2 & 1. \end{matrix}$$

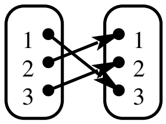
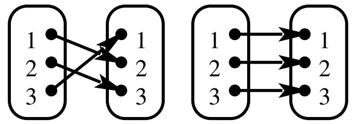
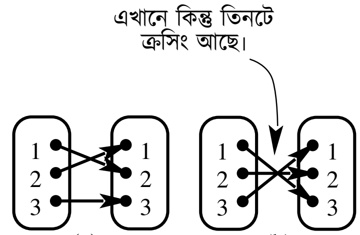


Fig 3



(a) (b)

Fig 4



(a) (b)

Fig 5

লক্ষ কর যে, determinant-টায় সত্যিই ঠিক 6-টা জিনিসই আছে! এই যে বললাম $1, 2, 3$-কে "সাজিয়ে লেখা", এটাকে আমরা একটা function বলে ভাবতে পারি। যেমন $3, 1, 2$ মানে হল Fig 3-এর function-টা। এখানে $f(1) = 3$ কারণ 1 নম্বর জায়গায় 3 আছে (অর্থাৎ 1 নম্বর row-এর 3 নম্বর entry-টা নেওয়া হয়েছে)। আবার 2 নম্বরে এসেছে 1, তাই $f(2) = 1$, ইত্যাদি। এরকম "সাজিয়ে লেখা"-র function-দের বলে একেকটা permutation. যদি এরকম সব permutation-এর set-টাকে S_3 বলি, তবে প্রত্যেকটা জিনিস হয়ে যাচ্ছে এরকম--

$$a_{1f(1)}a_{2f(2)}a_{3f(3)},$$

যেখানে $f \in S_3$. সুতরাং জিনিসগুলোকে সব বেশ গুছিয়ে লেখা গেল। এবার খালি একটাই কাজ বাকি। সেটা হল কারা যোগ হবে আর কারা বিয়োগ হবে, সেটা ঠিক করা। বিভিন্ন উদাহরণ নিয়েই Leibnitz দেখে নিয়েছিলেন যে, ঠিক অর্ধেক ক্ষেত্রে $+$ বসছে আর বাকি অর্ধেকে $-$. প্রশ্ন হল কোথায় কী বসবে, সেটা ধাঁ করে কী করে বলা যাবে? অনুমান করতে পারি যে, এই সমস্যাটা Leibnitz-কে বেশ খানিকটা ভুগিয়েছিল, যদিও শেষ পর্যন্ত তিনি এরও একটা চমৎকার সমাধান বার করেন। এবার সেটা বলি।

ধরো এই term-টা নিলাম, $a_{12}a_{23}a_{31}$. এখানে permutation-টা হল $f(1) = 2,\ f(2) = 3,\ f(3) = 1$. প্রথমে এর ছবিটা এঁকে নাও, তবে Fig 4(a) পাবে। এবার গুণে দ্যাখো তীর চিহ্নগুলো কতবার পরস্পরকে ক্রস করছে। ছবিটার দিকে তাকালেই বুঝবে যে আমাদের ক্ষেত্রে সংখ্যাটা হল 2. যেহেতু এটা একটা even (জোড়) সংখ্যা, তাই এখানে চিহ্নটা হবে '$+$'. যদি odd (বিজোড়) হত, তবে নিতাম '$-$'. খুবই অদ্ভুত নিয়ম সন্দেহ নেই। কায়দাটা সহজ হলেও মাথায় আসা মোটেই সহজ নয়। আরেকটা উদাহরণ নিই।

Example 1: $a_{11}a_{22}a_{33}$-এর বেলায় চিহ্নটা কী হবে?

SOLUTION: Fig 4(b)-র দিকে তাকালেই বুঝবে যে, এখানে কোনো ক্রসিংই নেই, মানে ক্রসিংএর সংখ্যা 0, যেটা একটা even সংখ্যা। হ্যাঁ, 0 একটা even সংখ্যা, কারণ 0-কে $0 = 2 \times 0$ হিসেবে লেখা যায়। অনেক ছাত্রছাত্রীর মনেই দেখেছি 0-টা even নাকি odd সে বিষয়ে সন্দেহ থাকে! যেহেতু 0 একটা even সংখ্যা, অতএব এখানে '$+$' চিহ্ন হবে। ■

Exercise 1: Fig 5-র দিকে এক ঝলক তাকিয়ে বল তো এই দুটো term-এর কী কী চিহ্ন হবে--

$$a_{12}a_{21}a_{33} \text{ এবং } a_{13}a_{22}a_{31}?$$

■

এইভাবে Leibnitz একটা নতুন ধারণার সৃষ্টি করলেন--**sign of a permutation**. যদি f একটা permutation হয়, তবে তার sign-টাকে $sgn(f)$ লেখা হয়। Fig 6-এ $n = 3$-এর জন্য যতরকম permutation হয়, তাদেরকে তীর চিহ্ন দিয়ে দেখানো হয়েছে। এদের দুটো row-তে ভাগ করেছি, প্রথম row-এর তিনজনের বেলায় চিহ্নটা হল $+1$, বাকিদের বেলায় -1.

Exercise 2: যদি $f(1) = 4,\ f(2) = 3,\ f(3) = 2$ এবং $f(4) = 1$ হয়, তবে f-এর sign কী হবে? ■

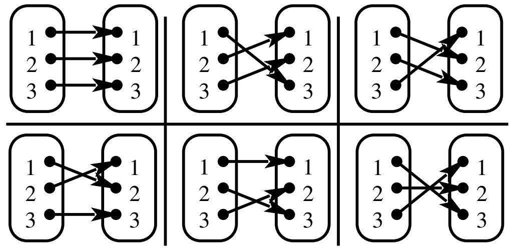


Fig 6

Exercise 3: ধরো $f : \{1, ..., n\} \to \{1, ..., n\}$ হল $f(i) = n - i + 1$. এটা একটা permutation. এর sign বার করো। ■

Exercise 4: ধরো $f : \{1, 2, 3\} \to \{1, 2, 3\}$ হল একটা permutation. যদি এটাকে extend করে $g : \{1, ..., 5\} \to \{1, ..., 5\}$ বানাই এইভাবে--

$$g(i) = \begin{cases} f(i) & \text{if } i = 1, 2, 3 \\ i & \text{otherwise} \end{cases},$$

তবে কি f আর g-এর sign একই হবে? ■

এবার আমরা determinant-এর প্রথম সংজ্ঞা লেখার জন্য তৈরী।

DEFINITION: **Determinant**

For a square matrix $A_{n\times n} \equiv ((a_{ij}))$ we define its **determinant** as

$$\sum_{f\in S_n} sgn(f) a_{1f(1)} a_{2f(2)} \cdots a_{nf(n)},$$

where S_n is the set of all permutations of $1, 2, ..., n$.

Leibnitz এটা নিয়ে এত খাটলেন বটে, কিন্তু শুনে দুঃখ হবে যে, কেউ সে আমলে ব্যাপারটাকে বড় একটা পাত্তা দেয় নি। তার কারণ জিনিসটার মধ্যে একটা ধাঁধার গন্ধ থাকলেও একটা প্রশ্ন থেকেই যায়--করে লাভটা কী? কোনো একগুচ্ছ equation-এর সমাধান আছে কিনা সেটা তো সরাসরি solve করার চেষ্টা করলেই বেরিয়ে যায়। বস্তুতঃ এই চিন্তাটা স্বয়ং Leibnitz-এরও মনে হয়ে থাকবে, কারণ তিনিও determinant নিয়ে তাঁর চিন্তাভাবনার অনেক কিছুই জীবদ্দশায় প্রকাশ করেন নি। পরে তাঁর কাগজপত্র ঘাঁটতে গিয়ে জানা গিয়েছে যে, determinant নিয়ে পরবর্তী কালের বিভিন্ন গণিতজ্ঞরা যা যা বার করেছেন, তার বেশ কয়েকটা Leibnitz-ও আগেই বার করেছিলেন, কিন্তু কাউকে সে কথা বলে যান নি। এইভাবে বহুদিন পর্যন্ত determinant ব্যাপারটা বলতে গেলে "পণ্ডিতদের খেলনা" হয়েই ছিল। অবসর সময়ে অনেকেই তাতে আলগোছে আঁচড় কেটেছেন, বিভিন্ন জিনিসপত্র প্রমাণও করেছেন, কিন্তু কেউই গুরুত্বটা অনুধাবন করতে পারেন নি। জিনিসটা যে সত্যই কাজের, সেটা প্রথম দেখিয়েছিলেন Gauss (গাউস)। সেই গল্পে এক্ষুণি আসছি। আগে দেখি Leibnitz-এর কাজের ভিত্তিতে কী কী করা যায়।

আমরা যেভাবে determinant-এর সংজ্ঞাটা লিখেছি, তাতে row ধরে ধরে এগিয়েছি। প্রত্যেকটা term-এ একটা করে permutation রয়েছে f, যেটা বলছে কোন row-এর জন্য কত নম্বর column-এর entry নেওয়া হয়েছে, যেমন

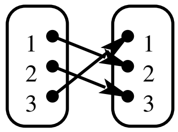


Fig 7

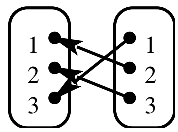


Fig 8

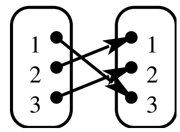


Fig 9

$a_{12}a_{23}a_{31}$-কে লিখতে পারি $a_{1f(1)}a_{2f(2)}a_{3f(3)}$ হিসেবে, যেখানে f-টা দেখিয়েছি Fig 7-এ। এখানে $f(1) = 2$ মানে প্রথম row-এর দ্বিতীয় entry-টা নেওয়া হয়েছে। চাইলে আমরা column ধরে ধরেও এগোতে পারতাম। মানে $a_{12}a_{23}a_{31}$-কে একটু আগেপরে করে সাজিয়ে লিখতে পারি $a_{31}a_{12}a_{23}$, এটার চেহারা হল $a_{g(1)1}a_{g(2)2}a_{g(3)3}$. এখানে g হল আগের f-টাকেই উল্টো করে দেখা। Fig 8 দেখলেই বুঝবে যে উল্টোনো মানে হল তীর চিহ্নগুলোর মুখ খালি ঘুরিয়ে দেওয়া। খালি তীর চিহ্নদের মুখ ঘুরিয়ে দিয়েছি বলেই তো আর খামোখা ক্রস করার সংখ্যা পাল্টে যাচ্ছে না! সুতরাং $sgn(g) = sgn(f)$ হতে বাধ্য। অতএব

$$sign(f)a_{1f(1)}a_{2f(2)}a_{3f(3)} = sign(g)a_{g(1)1}a_{g(2)2}a_{g(3)3}.$$

আমরা এই রকম $3! = 6$-টা term যোগ করেছিলাম, কারণ মোট ততগুলোই f আছে। f-দের চিহ্ন ওল্টালে ওই ছয়টাই ঘুরে ফিরে আসবে। যেমন ধরো Fig 8-এ যে function-টা দেখিয়েছিলাম, সেটা আসলে Fig 9-র function-টাই। সুতরাং আমরা যোগটা যাবতীয় f-এর উপরে না করে যাবতীয় g-এর উপরে করলেও একই ব্যাপার হবে। সব মিলিয়ে পেলাম এই theorem-টা--

THEOREM

For any square matrix $A_{n\times n} \equiv ((a_{ij}))$ we have

$$|A| = \sum_{g\in S_n} sign(g)a_{g(1)1}a_{g(2)2}\cdots a_{g(n)n},$$

where S_n is the set of all possible permutations of $1, 2, ..., n$.

যে যুক্তিতে এটা পেলাম তার অন্যতম সুবিধা হল determinant-দের বেলায় row নিয়ে যা যা করা যায়, column নিয়েও সে সবই করা যায়। এর বিভিন্ন নিদর্শন আমরা শীঘ্রই দেখব। আপাততঃ একটা সহজ প্রয়োগ হল এই theorem-টা--

THEOREM

For any square matrix A

$$|A| = |A'|.$$

এর প্রমাণটা নিতান্তই সহজ। ধরো A'-এর (i, j)-th entry-কে বলব b_{ij}. তার মানে $b_{ij} = a_{ji}$. এবার

$$|A'| = \sum_f sign(f)b_{1f(1)}b_{2f(2)}b_{3f(3)}.$$

এইটা সরাসরি সংজ্ঞা থেকেই লিখে দিয়েছি। এবার b-গুলোকে a দিয়ে লিখব, মানে যেখানেই b_{ij} দেখব সেগুলোকে a_{ji} করে দেব--

$$|A'| = \sum_f sign(f)a_{f(1)1}a_{f(2)2}a_{f(3)3}.$$

কিন্তু এটাই তো $|A|$-কে column ধরে ধরে লিখলে যা হয় সেটা। অতএব $|A'| = |A|$, ব্যস্ প্রমাণ শেষ!

DAY 34
বার করার নানা কায়দা

আগেই বলেছি যে determinant দিয়ে equation-এর সমাধানের অস্তিত্ব পরীক্ষা করা ছাড়া আর কিছু হয় বলে বহুদিন পর্যন্ত লোকের জানা ছিল না। সুতরাং Gauss-এর আগে পর্যন্ত এ বিষয়ে যা কাজ হয়েছে সব কিছুরই বক্তব্য একই--কী করে বড় বড় matrix-এর determinant সহজে বার করা যায়। বার করে কী হবে সে বিষয়ে কোনোই ধারণা তৈরী হয় নি, নিছক খেলাচ্ছলে অনেকগুলো কায়দা বার করেছিলেন বিভিন্ন পণ্ডিত। এদের মধ্যে অধিকাংশ নিয়েই আধুনিক কালে কেউ মাথা ঘামায় না। কিন্তু কয়েকটা বেশ দরকারী, সেগুলো নিয়ে আলোচনা করি।

34.1 Triangular/diagonal

কোনো triangular বা diagonal matrix-এর determinant বার করা খুব সোজা। Diagonal element-গুলোকে ধরে গুণ করে দিলেই হল।

Example 2: এই তিনটে matrix-এর determinant বার কর।

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 7 & 4 & 0 \\ 5 & 0 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

SOLUTION: এখানে A হল upper triangular, B হল lower triangular, আর C হল diagonal. কিন্তু সবার ক্ষেত্রেই diagonal entry-গুলো হল $1, 4, 6$. সুতরাং তিনজনেরই determinant হবে $1 \times 4 \times 6 = 24$. এমনটা কেন হচ্ছে, সেটা খুব সহজেই দেখা যায় Leibnitz-এর দেওয়া সংজ্ঞা থেকে। মনে কর একটা 5×5 upper triangular matrix আছে, এখানে $5! = 120$-টা term থাকবে। Fig 10 দ্যাখো। এখানে '$\times$' মানে যে entry-গুলো nonzero হতে পারে[1]। প্রতিটা term-ই হল পাঁচটা করে entry-র গুণফল। বলাই বাহুল্য যে, কোনো term-এর একটা entry-ও যদি শূন্য হয়, তবে term-টা শূন্য হয়ে যাবে। এবার প্রতি term-এই তো প্রথম column থেকে একটা entry থাকতে হবে। সেখানে তো শেষের 4-টে শূন্য, সুতরাং প্রথমটা নেওয়া ছাড়া পথ নেই। সুতরাং এই term-এ প্রথম row-টা ব্যবহার হয়ে গেল (Fig 11)। এবার দ্বিতীয় column থেকে একটা entry চাই। প্রথম row-টা ব্যবহার হয়ে গেছে, সুতরাং দ্বিতীয় row বা তার নীচে থেকে নিতে হবে। এখানেও খালি একটাই entry আছে যেটা nonzero হতে পারে (Fig 12)। এইভাবে দেখতে পাচ্ছ যে, খালি একটাই term পড়ে থাকছে, সেটা হল diagonal entry-গুলোর গুণফল--

$$a_{11}a_{22}\cdots a_{55}.$$

এখানে 1 নম্বর row-এর 1 নম্বর column, তারপর 2 নম্বর row-এর 2 নম্বর column, এইভাবে নেওয়া হয়েছে। সুতরাং permutation-টা হল Fig 13-এর মত। দেখতেই পাচ্ছ যে এখানে ক্রসিংএর সংখ্যা 0, যেটা একটা even সংখ্যা। তাই +

[1] শূন্য হলেও কোনো আপত্তি নেই।

Fig 10

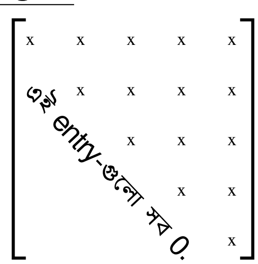


Fig 11

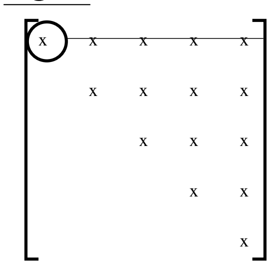


Fig 12

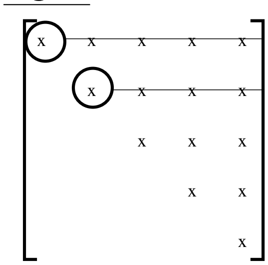
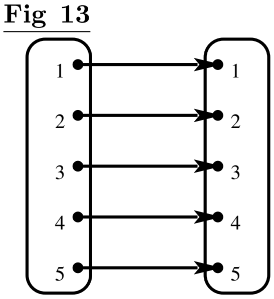


Fig 13



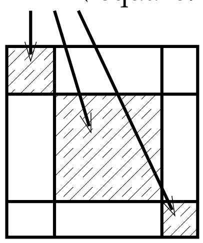


Fig 14

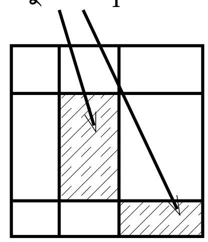


Fig 15

চিহ্ন নেব। ■

34.2 Block triangular matrix

আমরা প্রথম অধ্যায়ে block matrix বা partitioned matrix-এর কথা শিখেছিলাম, মনে আছে? যেমন

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

যেখানে

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix}, \\ C &= \begin{bmatrix} 13 & 14 & 15 \end{bmatrix}, \quad D = \begin{bmatrix} 16 \end{bmatrix}. \end{aligned}$$

আমরা এখানে square matrix নিয়ে কাজ করব, এবং diagonal block-গুলোকেও নেব square. মানে Fig 14-এর মত, Fig 15-এর মত নয়।

এবার যদি off-diagonal block-গুলো (মানে diagonal ছাড়া বাকি block-গুলো) সবাই zero matrix হয় (কোনো diagonal block-ও শূন্য হলে আপত্তি নেই), তবে পুরো matrix-টাকে বলব block diagonal. যেমন--

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix} \text{ বা } \begin{bmatrix} A & O & O \\ O & B & O \\ O & O & F \end{bmatrix}.$$

একইভাবে block triangular matrix-রা হল এইরকম দেখতে--

$$\begin{bmatrix} A & B \\ O & C \end{bmatrix} \text{ বা } \begin{bmatrix} A & O & O \\ B & C & O \\ D & E & F \end{bmatrix}.$$

এরকম matrix-দের determinant বার করার সহজ কায়দা হল diagonal block-গুলোর determinant বার করে গুণ করে দেওয়া, যেমন--

$$\begin{vmatrix} A & B \\ O & C \end{vmatrix} = |A| \times |C| \text{ আর } \begin{vmatrix} A & O & O \\ B & C & O \\ D & E & F \end{vmatrix} = |A| \times |C| \times |F|.$$

এমনটা কেন হয় সেটা Leibnitz-এর দেওয়া সংজ্ঞা থেকেই বোঝা যায়। যদিও আমরা এখানে তার প্রমাণের মধ্যে যাচ্ছি না। বরং একটা প্রয়োগ দেখা যাক।

Example 3: $\begin{vmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 1 & 1 & 9 & 0 & 0 \\ 2 & 3 & 1 & 5 & 6 \\ 0 & 2 & 8 & 1 & 2 \end{vmatrix} = ?$

SOLUTION: এটা একটা block triangular matrix. সেটা বুঝবে এইভাবে ভাবলেই--

$$\left[\begin{array}{cc|c|cc} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ \hline 1 & 1 & 9 & 0 & 0 \\ \hline 2 & 3 & 1 & 5 & 6 \\ 0 & 2 & 8 & 1 & 2 \end{array}\right]$$

সুতরাং determinant-টা হবে--

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \times 9 \times \begin{vmatrix} 5 & 6 \\ 1 & 2 \end{vmatrix} = -2 \times 9 \times 4 = -72.$$

■

Exercise 5: $\begin{vmatrix} 2 & 4 & 1 & 7 & 3 \\ 2 & 4 & 6 & 0 & 8 \\ 0 & 0 & 9 & 3 & 0 \\ 0 & 0 & 8 & 5 & 6 \\ 0 & 0 & 0 & 0 & 2 \end{vmatrix} = ?$ ■

34.3 Van der Monde matrix

Van der Monde (ভ্যান্ ডার মণ্ডে) matrix হল একটা বিশেষ ধরণের square matrix, যার সংজ্ঞাটা এইরকম--

DEFINITION: **Van der Monde matrix**

A **Van der Monde** matrix of order n is a matrix of the form

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix},$$

where $a_1, ..., a_n$ are any scalars.

যেমন একটা উদাহরণ হল $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 4 & 9 & 16 \end{bmatrix}$. এই জাতীয় matrix অংকের বিভিন্ন শাখায় উদয় হয়। এদের নিয়ে বিস্তারিত আলোচনা এখানে সম্ভব নয়। আমাদের খালি বলব এর determinant-টা কী হয়।

THEOREM

The determinant of the Van der Monde matrix,

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix},$$

is

$$\prod_{i<j}(a_i - a_j).$$

এর প্রমাণ আমরা এ বইতে করব না। খালি theorem-টার বক্তব্যটা হজম করে রাখব। একটা $n \times n$ সাইজে Van der Monde matrix বানাতে n-খানা সংখ্যা লাগে, $a_1, ..., a_n$ (সবাই আলাদা আলাদা নাও হতে পারে)। এবার determinant বানানোর জন্য সংখ্যাগুলোকে পরপর লিখে নিতে হয়। তারপর প্রত্যেকটা সংখ্যা থেকে তার পরের সবগুলো সংখ্যা বিয়োগ করে দিতে হয়। এইভাবে যতগুলো বিয়োগফল পাওয়া গেল, সবগুলোকে গুণ করে দিলেই determinant পাওয়া যায়। একটা উদাহরণ না দেখালেই নয়!

Example 4: $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1^2 & 2^2 & 3^2 & 4^2 \\ 1^3 & 2^3 & 3^3 & 4^3 \end{vmatrix} = ?$

SOLUTION: এখানে $a_1 = 1,\ a_2 = 2,\ a_3 = 3$ এবং $a_4 = 4$. এদেরকে পরপর লিখে ফেলি--

□,□,□,□.

1-এর পরে আছে $2, 3, 4$. এদেরকে 1 থেকে বিয়োগ করে পাচ্ছি $1 - 2 = -1,\ 1 - 3 = -2$ আর $1 - 4 = -3$.
2-এর পরে আছে $3, 4$. এদেরকে 2 থেকে বিয়োগ করে পাচ্ছি $2 - 3 = -1$ আর $2 - 4 = -2$.
3-এর পরে আছে শুধু 4. একে 3 থেকে বিয়োগ করে পাচ্ছি $3 - 4 = -1$.
সবগুলো বিয়োগফলকে একসঙ্গে গুণ করে দিলে হচ্ছে--

$$-1 \times (-2) \times (-3) \times (-1) \times (-2) \times (-1) = 12.$$

সেটাই হবে determinant-টা। ■

Exercise 6: $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1^2 & 2^2 & 4^2 \end{vmatrix} = ?$ ■

34.4 Laplace expansion

এবার যে কায়দাটা বলব, সেটাও আসলে Leibnitz-এরই আবিষ্কার। কিন্তু তিনি সেটা প্রকাশ করে যাননি। কায়দাটা নতুন করে বার করেন Laplace (লাপ্লাস), তাই তাঁর নামটাই চালু হয়েছে। পরে Leibnitz-এর অপ্রকাশিত কাগজপত্র ঘাঁটতে গিয়ে জানা গেছে যে, তিনি বহু আগেই একই জিনিস বার করেছিলেন।

কায়দাটা হল ধাপে ধাপে determinant বার করা। যদি কোনো 1×1 সাইজের matrix দিই, তবে সেটা বস্তুতঃ একটামাত্র সংখ্যা। এরকম matrix-এর determinant হল সেই সংখ্যাটা নিজেই। যদি কোনো 2×2 সাইজের matrix দিই, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, তবে তার determinant হবে $ad-bc$. এ দুটো ধাপ মনে রাখা নেহাতই সহজ। এবার ধরো তোমাকে বললাম এই matrix-টার determinant বার করতে--

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Leibnitz-এর কায়দায় এগোলে এর determinant আমরা জানি--

$$|A| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}.$$

এখানে প্রত্যেকটা term-এ প্রত্যেকটা row আর column থেকে ঠিক একটা করেই entry নেওয়া হয়েছে। কিন্তু term-গুলোকে মনে রাখা সহজ নয়। তাই হায়ার সেকেণ্ডারীর সময়ে লোকে সাধারণতঃ একটা কৌশল করে এটা মনে রাখে। যেকোনো একটা row (বা column) মনে মনে ঠিক করে নাও, ধরো প্রথম row-টা। এটা হল $\begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix}$. যেখানে যেখানে এই তিনজন আছে, সেখান থেকে তাদের 'কমন' নিয়ে determinant-টাকে এভাবে লেখা যায়--

$$|A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

এর মধ্যে যে determinant-গুলো আছে, সেগুলো সব 2×2 সাইজের matrix-দের, তাই চোখে দেখেই কষে দেওয়া যায়। Determinant-এর এই রূপটা হায়ার সেকণ্ডারী থেকেই আমাদের চেনা। এই কায়দাটার সুবিধা হল চিহ্ন নিয়ে অত মাথা ঘামাতে হয় না। এখানে আমরা একটা 3×3 সাইজের matrix-এর determinant-কে কয়েকটা 2×2 সাইজের matrix-এর determinant দিয়ে লিখতে পারলাম। এই কাজটা আসলে যেকোনো $n\times n$ সাইজের matrix-এর জন্যই করা যায় (যেখানে $n\geq 2$)। মানে এরকমভাবে একটা row বা column ধরে n-খানা $(n-1)\times(n-1)$ সাইজের matrix-এর determinant-এ ভেঙে লেখা যায়। এটাকেই বলে **Laplace expansion**, যেটা এবার আলোচনা করব।

প্রথমে যেকোনো একটা row বা column ঠিক করে নিতে হয়। যেমন এখানে আমরা প্রথম row নিয়ে কাজ করেছি। এবার সেই row-এর একেকটা করে entry নাও, সেই entry-র row আর column কেটে দাও , যেমন $n=3$-এর জন্য a_{11}-র বেলায় দেখিয়েছি Fig 16-এ। Matrix-টার যে অংশটা পড়ে রইল সেটা আবার একটা square matrix-ই হবে। তার determinant বার করতে হবে। একে বলে একটা **minor**. যেহেতু আমরা a_{11} নিয়ে কাজ শুরু করেছিলাম, তাই একে বলব $(1,1)$-th minor, আর লেখার সময়ে লিখব A_{11}. একইভাবে প্রথম row-এর বাকি entry-গুলোর জন্যও করব। Fig

Fig 16

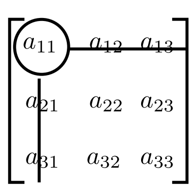
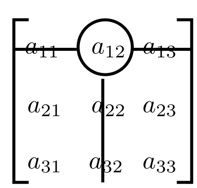


Fig 17



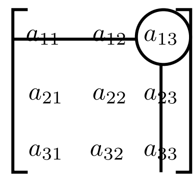


Fig 18

17-এ দেখিয়েছি a_{12}-র কেসটা, আর Fig 18-এ a_{13}-র টা। এবার প্রথম row-এর entry-গুলোকে তাদের minor-গুলো দিয়ে গুণ করে দাও। এর ফলে general n-এর জন্য n-খানা গুণফল পাবে--

$$a_{11}A_{11},\quad a_{12}A_{12},\quad a_{13}A_{13},\quad \ldots,\quad a_{1n}A_{1n}.$$

এবার প্রত্যেকের আগে একটা করে চিহ্ন বসাতে হবে। নিয়মটা খুবই সহজ--প্রত্যেকটা গুণফলই দেখতে $a_{ij}A_{ij}$-এর মত। তোমাকে খালি দেখে নিতে হবে $i+j$-টা even (জোড়) নাকি odd (বিজোড়)। যদি even হয় (যেমন $a_{11}A_{11}$-র বেলায় $1+1=2$ হল even), তবে চিহ্নটা হবে '$+$', আর যদি odd হয় (যেমন $a_{12}A_{12}$-র বেলায় $1+2=3$ হল odd), তবে '$-$'। সুতরাং সব মিলিয়ে হবে

$$|A| = a_{11}A_{11} - a_{12}A_{12} + \cdots + (-1)^{1+n}a_{1n}A_{1n}.$$

এই ফর্মুলাটার সুবিধা হল চিহ্ন ঠিক করা সহজ। আমরা চাইলে অন্য কোনো row নিয়েও এগোতে পারতাম। যদি i-th row ধরে এগোতাম, তবে পেতাম--

Laplace expansion (row version)

Let $A_{n\times n}$ have entries a_{ij}. Let A_{ij} be the determinant of the submatrix obtained by deleting the i-th row and j-th column of A. Then

$$|A| = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}A_{ij}.$$

This is called **Laplace expansion along the i-th row**.

যেহেতু transpose নিলে determinant বদলায় না, অতএব আমরা কোনো column বরাবরও কাজটা করতে পারি। যেমন, j-th column বরাবর করলে--

Laplace expansion (column version)

Let $A_{n\times n}$ have entries a_{ij}. Let A_{ij} be the determinant of the submatrix obtained by deleting the i-th row and j-th column of A. Then

$$|A| = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}A_{ij}.$$

This is called **Laplace expansion along the j-th column**.

দুইক্ষেত্রেই ওই $(-1)^{i+j}$ অংশটা লক্ষ কর--ওটাই চিহ্নটাকে নিয়ন্ত্রণ করছে। চিহ্নটাকে A_{ij}-র মধ্যে ঢুকিয়ে নিলে লিখতে সুবিধা হয়, মানে $(-1)^{i+j}A_{ij}$-র একটা আলাদা নাম দিয়ে নিলে। একে বলে (i,j)-th **cofactor**. সংজ্ঞাটা লিখে রাখি--

DEFINITION: **Cofactor**

Let A be any square matrix. Let A_{ij} be the determinant of the submatrix obtained by deleting the i-th row and j-th column of A. Then

$$C_{ij} = (-1)^{i+j} A_{ij}$$

is called the (i,j)-th **cofactor** of A.

Cofactor দিয়ে লিখলে Laplace expansion আরও সহজে লেখা যায়--

$$\sum_i a_{ij}C_{ij} \text{ বা } \sum_j a_{ij}C_{ij}.$$

34.4.1 উদাহরণ

Laplace expansion দিয়ে একটা $n \times n$ সাইজের matrix-এর determinant-কে n-খানা $(n-1)\times(n-1)$ সাইজের matrix-এর determinant দিয়ে লেখা যায়। সেই $(n-1)\times(n-1)$ সাইজের matrix-গুলোর উপর আবার Laplace expansion লাগিয়ে তাদের determinant-দের $(n-2)\times(n-2)$ সাইজের matrix-এর determinant দিয়ে লেখা যায়। এইভাবে যতই Laplace expansion লাগাবে ততই matrix-গুলোর সাইজ ছোটো হয়ে আসবে, কিন্তু ওদের সংখ্যা বৃদ্ধি পাবে। যদি 3×3 বা 4×4 জাতীয় ছোটোখাটো matrix নিয়ে কাজ করতে হয়, তবে এভাবে এগোনো পোষায়, নইলে অসুবিধা। নীচের অংকটায় আমরা কাজ করব একটা 4×4 সাইজের matrix নিয়ে, দুবার Laplace expansion লাগিয়ে 2×2-তে এনে ফেলা যাবে, এবং সেগুলো চোখে দেখেই বার করে ফেলা যাবে।

Example 5: $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & 0 \\ 4 & 5 & 6 \end{vmatrix} = 0?$

SOLUTION: এখানে দ্বিতীয় row-তে অনেকগুলো শূন্য আছে। তাই সেই row বরাবর Laplace expansion লাগালে পরিশ্রম কম হবে। Determinant-টা হবে--

$$-0\times\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} + 5\times\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} - 0\times\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}.$$

এখানে প্রথম আর শেষ determinant দুটো কষ্ট করে আর বার করতে হবে না, কারণ ওদের সঙ্গে শূন্য গুণ হয়ে আছে। তাই উত্তর হবে--

$$5\times\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = 5\times(-6) = -30.$$

■

Exercise 7: কোনো triangular matrix-এর determinant যে তার diagonal entry-গুলো গুণ করে দিলেই বার হয়ে যায়, সেটার প্রমাণ আসে Laplace expansion থেকেই। প্রমাণটা কঠিন নয় মোটেই। চেষ্টা করে দেখো।
HINT:

প্রথমে একটা উদাহরণ নিয়ে ভাবো। ধরো matrix-টা হল $\begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}$. প্রথম row বরাবর এর determinant-এর Laplace expansion করে শুরু করো। ■

Exercise 8: Expanding by Laplace's method, show that

$$\begin{vmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{vmatrix} = (a^2+b^2+c^2+d^2)^2.$$

■

যদিও বড় বড় matrix-দের determinant বার করার জন্য Laplace expansion খুব সুবিধাজনক নয়, কিন্তু এর একটা ব্যতিক্রম আছে। যদি matrix-টার কোনো একটা বিশেষ প্যাটার্ণ থাকে তবে একবার Laplace expansion লাগালেই অনেক সময়ে সেটা একই প্যাটার্ণের ছোটো matrix-এ পরিণত হয়। যদি D_n হয় $n\times n$ সাইজের matrix-টার determinant,তবে এইভাবে D_n-কে D_{n-1} দিয়ে প্রকাশ করা যায়। এইবার সেই একই কায়দায় D_{n-1}-কে D_{n-2} দিয়ে, তাকে আবার D_{n-3} দিয়ে, এইভাবে D_2 বা D_1-এ পৌঁছে গেলে সেগুলো চোখে দেখেই বার করা যায়। নীচের অংকটা এরকম একটা উদাহরণ।

Exercise 9: Let

$$D_n = \begin{vmatrix} 1 & x & 0 & 0 & \cdots & 0 & 0 \\ x & 1 & x & 0 & \cdots & 0 & 0 \\ 0 & x & 1 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & x & 1 \end{vmatrix}$$

be a determinant of order n. Prove that

$$D_4 - D_3 + x^2 D_2 = 0.$$

Hence prove that

$$\begin{vmatrix} 1 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 1 \end{vmatrix} = \frac{5}{16}.$$

■

এইরকম আরেকটা অংক তোমার হাত পাকানোর জন্য দিই।

Exercise 10: Let

$$D_n = \begin{vmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & \cdots & a_2 & 0 \\ \vdots & \vdots & \cdot^{\cdot^{\cdot}} & \vdots & \vdots \\ 0 & a_{n-1} & \cdots 0 & 0 & 0 \\ a_n & 0 & \cdots 0 & 0 & 0 \end{vmatrix}.$$

Show that $D_n = (-1)^{n+1} D_{n-1}$. Hence find a formula for D_n. ■

34.5 RREF

RREF-এর কথা মনে আছে নিশ্চয়ই? সেই সর্বরোগের অব্যর্থ দাওয়াই? বলাই বাহুল্য যে, determinant বার করতেও RREF ব্যবহার করা যায়। বড় বড় matrix-এর determinant বার করার জন্য এটাই ভালো পথ (বিশেষ করে কম্পিউটারের সাহায্যে)। কায়দাটা হল matrix-টার RREF বার করা। যদি শেষ পর্যন্ত matrix-টা identity matrix-এ পরিণত না হয়, তবে determinant অবশ্যই 0 হবে। যদি identity matrix হয়, তবে দেখতে হবে কতবার scale আর swap করতে হয়েছে। প্রত্যেকটা scale করার ধাপে যা যা দিয়ে ভাগ করেছ, সেইগুলোকে মনে রাখতে হবে। ধরো তাদের গুণফল হল p. যদি swap-এর সংখ্যা even হয়, তবে এই p-টাই হবে determinant. আর যদি swap-এর সংখ্যা odd হয়, তবে determinant হবে $-p$. একটা উদাহরণ দেখাচ্ছি।

Example 6: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 8 & 10 \end{bmatrix}$-এর determinant বার করো RREF দিয়ে।

SOLUTION: Gauss-Jordan elimination দিয়ে RREF বার করার ধাপগুলো দেখাচ্ছি।

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 8 & 10 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 0 & -6 & -11 \end{bmatrix} \xrightarrow{Swap, \div(-6)} \begin{bmatrix} 1 & 0 & -\frac{2}{3} \\ 0 & 1 & \frac{11}{6} \\ 0 & 0 & -6 \end{bmatrix} \xrightarrow{\div(-6)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

প্রত্যেক ধাপে খালি দুটো জিনিস খেয়াল রাখছি, swap লাগছে কিনা, এবং scale করা হলে কী দিয়ে ভাগ হচ্ছে। শেষ পর্যন্ত identity matrix এসেছে। ভাগ হয়েছে দুবার, প্রত্যেকবারই -6 দিয়ে। এদেরকে গুণ করলে হয় $(-6) \times (-6) = 36$. যেহেতু বিজোড়-সংখ্যকবার swap করা হয়েছে, তাই একটা মাইনাস চিহ্ন আসবে। ফলে determinant হবে -36. ■

DAY 35
Properties

Determinant কাকে বলে এবং কী করে বার করে, সেটা আমরা শিখেছি। আজকে আমরা determinant-এর বিভিন্ন ধর্ম আলোচনা করব।

35.1 Determinant of product

ধরো A আর B দুটো একই সাইজের square matrix. তা হলে ওদের দুজনের determinant আলাদা করে বার করে গুণ করে দিলেই AB-র determinant পেয়ে যাবে! অংকের ভাষায়--

THEOREM

For any two square matrices $A_{n\times n}$ and $B_{n\times n}$

$$|AB| = |A| \cdot |B|.$$

দুটো square matrix গুণ করলে determinant-ও গুণ হয়ে যায় বটে, কিন্তু তা বলে যেসব ছাত্র ফূর্তির চোটে বলে বসে যে, দুটো matrix যোগ বা বিয়োগ করলে determinant-ও যোগ-বিয়োগ হয়ে যায়, তারা ভয়ানক মিথ্যাবাদী, তাদের কথায় মোটেই কান দিও না! এরকম কুচিন্তার প্রতিষেধক হিসেবে আগেভাগেই নীচের অংকটা করে রাখো।

আমরা ক্বচিৎ কখনো লেপ কম্বল কিম্বা তোষকবালিশ এ-সব একটু-আধটু খাই বটে, কিন্তু যারা বলে আমরা খাট পালং কিম্বা টেবিল চেয়ার খাই, তারা ভয়ানক মিথ্যাবাদী।

--হ য ব র ল

Exercise 11: এমন দুটো matrix দাও A, B, যাতে

$$|A+B| \neq |A|+|B|, \quad |A-B| \neq |A|-|B| \text{ এবং } |-A| \neq -|A|$$

হয়। ■

35.2 Invertibility এবং rank-এর সাথে determinant-এর সম্পর্ক

একটা matrix-এর invertible (বা nonsingular) হওয়ার সঙ্গে determinant-এর ওতপ্রোত সম্পর্ক আছে। তার মধ্যে সবচেয়ে গুরুত্বপূর্ণ হল এইটা--যদি A একটা $n \times n$ সাইজের matrix হয়, তবে A-টা nonsingular হওয়া এবং $|A| \neq 0$ হওয়া সমার্থক।

THEOREM

A is an invertible matrix $\iff |A| \neq 0$. In this case, $|A^{-1}| = 1/|A|$.

এর প্রমাণ অতি সহজ। ধরো A হল invertible. তার মানে A^{-1} আছে, এবং $AA^{-1} = I$. এবার দুদিকের determinant নেব। যেহেতু গুণের বেলায় transpose-কে ভেঙে লেখা যায়, তাই

$$|A| \cdot |A^{-1}| = |I| = 1.$$

সুতরাং $|A| \neq 0$ হতে বাধ্য, এবং $|A^{-1}| = \frac{1}{|A|}$. এইটা ব্যবহার করে একগুচ্ছ vector মিলে $\mathbb{R}^n$-এর basis হচ্ছে কিনা চট্ করে পরীক্ষা করা যায়, কারণ যদি $\{\vec{v}_1, ..., \vec{v}_n\} \subseteq \mathbb{R}^n$ নাও, আর তাদেরকে পাশাপাশি column-এর মত দাঁড় করিয়ে একটা matrix বানাও, তবে সেই matrix-টার nonsingular হওয়া আর $\{\vec{v}_1, ..., \vec{v}_n\}$-টা একটা basis হওয়া সমার্থক। নীচের অংকে এর একটা প্রয়োগ দিলাম।

Example 7: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). Values of K for which S is linearly dependent in $\mathbb{R}^3$, where $S = \{(1,2,1),(K,3,1),(2,K,0)\}$ are

(A) $-2, -1$ (B) $-2, 1$ (C) $2, -1$ (D) $2, 1$.

(2019.CC6.1j)

SOLUTION:এখানে vector-গুলোকে একটার উপর একটা লিখে পাবে এই matrix-টা--

$$\begin{bmatrix} 1 & 2 & 1 \\ K & 3 & 1 \\ 2 & K & 0 \end{bmatrix}.$$

আমরা অবশ্য একটু আগে ওদেরকে column-এর মত পাশাপাশি দাঁড় করিয়ে matrix বানাবার কথা বলেছিলাম। সুতরাং এই matrix-টা হল সেইটার transpose. কিন্তু transpose নিলে যেহেতু determinant বদলায় না, তাই আমাদের যুক্তি এখানেও একইভাবে খাটবে। একটু কষলেই দেখবে এর determinant আসছে $K^2 - K - 2 = (K-2)(K+1)$. সুতরাং উত্তর যে (C) হবে বুঝতেই পারছ। ■

এতক্ষণ যে আলোচনা করলাম সেটা খালি square matrix-দের বেলায় খাটে। যদি square না হয়, তাহলেও determinant-এর সঙ্গে rank-এর একটা যোগ আছে। সেখানে অবশ্য পুরো matrix-টার determinant বার করা যায় না, খালি square আকারের submatrix-দের determinant নিয়ে কাজ করতে হয়।

THEOREM

Let $r(A_{m\times n}) = r$. Then there is an $r\times r$ submatrix P of A such that $|P| \neq 0$. Also any square submatrix $Q_{k\times k}$ for $k > r$ must have $|Q| = 0$.

এর মানে হল যদি $r(A) = r$ হয়, তবে A-র মধ্যে এমন অন্ততঃ একটা $r\times r$ সাইজের submatrix পাবে, যেটার determinant শূন্য নয়। যদি তুমি এর চেয়ে বড় কোনো সাইজের submatrix নাও A-র মধ্যে, তবে তার determinant শূন্য হতে বাধ্য। একটা উদাহরণ দেখে হজম করে নেওয়া যাক।

Example 8: এই matrix-টার rank কত?

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 10 & 20 & 30 & 50 \\ 11 & 22 & 33 & 54 \end{bmatrix}.$$

এর মধ্যে rank-এর সমান সাইজের একটা submatrix দেখাও, যার determinant শূন্য নয়।

SOLUTION: একটু খুঁটিয়ে দেখলেই লক্ষ্য করবে যে, তৃতীয় row-টা প্রথম দুটোর যোগফল। সুতরাং rank হতে পারে বড়জোর 2. এদিকে প্রথম row দুটো মোটেই পরস্পরের multiple নয়। তার মানে প্রথম দুটো row মিলে row space-এর একটা basis হচ্ছে। তাই rank হল 2. এবার 2×2 সাইজের submatrix-এর খোঁজ করা যাক। আমরা এক্ষুণি দেখলাম যে, প্রথম দুটো row মিলে row space-এর একটা basis হয়। আমরা খালি ওই দুটো row-কেই রাখব। এর ফলে একটা submatrix পেলাম--

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 10 & 20 & 30 & 50 \end{bmatrix}.$$

কিন্তু এটা এখনো 2×2 সাইজের নয়। যেহেতু এর row-গুলো linearly independent আছেই, তাই এই submatrix-টার rank-ও সেই 2-ই আছে। সুতরাং আমরা এমন দুটো column পাব, যারা মিলে column space-এর basis হবে। এখানে প্রথম আর শেষ column নিলেই হবে। আমরা খালি ওই column দুটোই রাখব। ব্যস্‌ আমাদের submatrix তৈরী--

$$\begin{bmatrix} 1 & 4 \\ 10 & 50 \end{bmatrix}.$$

যেহেতু এর column দুটো linearly independent (সেভাবেই column নির্বাচন করলাম), তাই এটার rank হল 2, আর সাইজও 2×2, তাই একটা nonsingular হতে বাধ্য। সুতরাং এর determinant শূন্য হতে পারে না। ■

এই যুক্তিটা যেকোনো $A_{m\times n}$-এর জন্যই খাটবে। যদি $r(A) = r$ হয়, তবে তুমি প্রথমে এমন r-খানা row পাবে, যারা মিলে $\mathscr{R}(A)$-র basis হয়। খালি সেই r-খানা row-ই রাখলে একটা submatrix পাবে $B_{r\times n}$. এটারও RREF হবে

r. সুতরাং এমন r-খানা column পাবে, যারা $\mathscr{C}(B)$-এর basis হয়। খালি সেই column-গুলোকেই রাখলে পাবে $P_{r\times r}$, যার determinant অবশ্যই nonzero হবে।

Exercise 12: ধরো $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$. এর rank কত? সেই সাইজের একটা submatrix বার করো, যার determinant শূন্য নয়। ■

আমাদের theorem-এর প্রথম অংশের প্রমাণের কায়দাটা তাহলে বুঝলে। দ্বিতীয় অংশের প্রমাণটা আরো সংক্ষিপ্ত। দ্বিতীয় অংশে বলা ছিল যে, যদি $r(A) = r$ হয়, এবং $Q_{k\times k}$ তার একটা submatrix হয়, যেখানে $k > r$, তবে $|Q| = 0$ হতে বাধ্য। কারণ A-র rank-ই মোটে r. আর Q হল গিয়ে সেই A-র submatrix, সুতরাং তার rank তো r-কে ছাড়িয়ে যেতে পারে না। কিন্তু $k > r$. অতএব $r(Q_{k\times k}) < k$. তার মানে Q হল singular matrix. সুতরাং $|Q| = 0$.

35.3 Block matrix

এবার একটা অদ্ভুত দেখতে ধর্মের কথা বলব, যার উপযোগিতা প্রথম দর্শনে ঠিক বোঝা যায় না।

> **THEOREM**
>
> If $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is a square matrix with A nonsingular, then
>
> $$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \cdot |D - CA^{-1}B|.$$

একটা উদাহরণ না নিলে বোঝা শক্ত এখানে হচ্ছেটা কী!

Example 9: ধরো তোমাকে এই matrix-টার determinant বার করতে বললাম।

$$\begin{bmatrix} 5 & 2 & 3 & 4 \\ 7 & 3 & 7 & 8 \\ 0 & 3 & 4 & 2 \\ 9 & 3 & 2 & 1 \end{bmatrix}.$$

ভেবে বলো, কী করে এই theorem-টা কাজে লাগাতে পারো।

SOLUTION: এটাকে আমরা প্রথমে এইভাবে partition করে নেব--

$$\left[\begin{array}{cc|cc} 5 & 2 & 3 & 4 \\ 7 & 3 & 7 & 8 \\ \hline 0 & 3 & 4 & 2 \\ 9 & 3 & 2 & 1 \end{array}\right].$$

একে যদি $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ বলি, তবে

$$A = \begin{bmatrix} 5 & 2 \\ 7 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 4 \\ 7 & 8 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 3 \\ 9 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}.$$

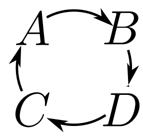


Fig 19

এখানে A-র determinant হল $|A| = 5 \times 3 - 2 \times 7 = 1 \neq 0$. সুতরাং A হচ্ছে nonsingular. অতএব আমাদের theorem-টা লাগানো যাবে। তার জন্য $D - CA^{-1}B$ বার করে রাখি। Fig 19 দেখলে এই ফর্মুলাটা খেয়াল রাখা সহজ। D থেকে শুরু করে ঘড়ির কাঁটার দিকে ঘুরতে হবে। এই প্রসঙ্গে একটা ভাষা শিখে রাখো। এই $D - CA^{-1}B$-কে বলে পুরো matrix-টার মধ্যে A-র **Schür complement** (শুর্ কম্প্লিমেন্ট)। এখানে এটা হল

$$D - CA^{-1}B = \cdots = \begin{bmatrix} -38 & -34 \\ 5 & 1 \end{bmatrix}$$

এর determinant হল $-38 \times 1 - 5 \times (-34) = 132$. সুতরাং পুরো matrix-টার determinant হবে

$$|A| \cdot |D - CA^{-1}B| = 1 \times 132 = 132.$$

■

এই উদাহরণে আমরা একটা 4×4 সাইজের matrix-এর determinant বার করলাম, কেবলমাত্র কয়েকটা 2×2 সাইজের matrix নিয়ে কাজ করে। কিন্তু এই theorem-এর প্রধান প্রয়োগ এটাই নয়। এর আসল প্রয়োগ হল যেখানে A-র সাইজ 1×1. সেক্ষেত্রে পুরো matrix-টা যদি $n \times n$ হয়, তবে তবে A-র Schür complement-এর সাইজ হবে $(n-1) \times (n-1)$. সুতরাং এই theorem-টা দিয়ে একটা $n \times n$ সাইজের determinant-কে একটা $(n-1) \times (n-1)$ সাইজের determinant দিয়ে লেখা যাচ্ছে। অনেক সময়ে যখন determinant-এর কোনো theorem আমরা induction দিয়ে প্রমাণ করি, তখন এটা খুব কাজে লাগে। তবে এই বইতে আমরা সেরকম প্রয়োগের মধ্যে যাব না।

35.4 Matrix-এর পরিবর্তনের সঙ্গে determinant-এর পরিবর্তন

ধরো একটা matrix-এর determinant তুমি কষ্ট করে বার করেছ। এবার যদি matrix-টাকে সামান্য বদলে দিয়ে ফের determinant বার করতে বলি, তবে একটা পথ হল একেবারে নতুন করে শুরু করা। কিন্তু অনেক সময়ে আগের determinant-টাকেই সামান্য বদলে নতুন determinant-টা পাওয়া সম্ভব। এবার সেরকম কিছু ক্ষেত্রের কথা বলব।

35.4.1 Swap করলে

যদি একটা square matrix-এর দুটো row-এর স্থানবিনিময় করো, তবে determinant-টার চিহ্ন ওল্টাবে মাত্র, আর কোনো পরিবর্তন হবে না। যেমন, আগে যদি determinant-টা 3 হয়ে থাকে, তবে এখন হয়ে যাবে -3. যদি আগে থাকত -5, তবে পরে হবে 5. একইরকম কথা বলা যায় দুটো column-এর স্থানবিনিময়ের বেলাতেও।

Example 10: যদি $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}$, $B = \begin{bmatrix} 7 & 8 & 10 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}$ আর $C = \begin{bmatrix} 2 & 1 & 3 \\ 5 & 4 & 6 \\ 8 & 7 & 10 \end{bmatrix}$ হয়, এবং $|A| = a$ হয়, তবে $|B|, |C|$ বার করো।

SOLUTION: এখানে B-কে পেয়েছি A-র প্রথম আর তৃতীয় row-এর স্থানবিনিময় করে, তাই $|B| = -|A| = -a$. একইভাবে, C-কে পাওয়া গেছে A-র প্রথম column দুটোর স্থানবিনিময় করে। তাই $|C| = -|A| = -a$. ■

Exercise 13: যদি $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 10 \\ 1 & 2 & 3 \end{bmatrix}$ হয়, এবং $|A| = a$ হয়, তবে $|B|$ বার করো।
■

Exercise 14: দেখাও যে, কোনো square matrix-এর দুটো row যদি একই হয়, তবে তার determinant শূন্য হবেই।
HINT: ওই row দুটোর স্থানবিনিময় করে দেখো, কী হয়।
■

এই জিনিসটা বেশ কাজে লাগে। তাই গুছিয়ে লিখে রাখি--

> THEOREM
>
> If a matrix A has two identical rows (or columns) then $|A| = 0$.

Exercise 15: আচ্ছা, ধরো একটা matrix আছে যার একটা row আর একটা column সমান, মানে row-টাকে খাড়া করে দিলেই column-টা হয়ে যায়। তাহলেও কি determinant-টা 0 হবে? ■

35.4.2 Scalar দিয়ে গুণ করলে

ধরো A হল একটা $n \times n$ সাইজের matrix, আর c হল একটা scalar. যদি c দিয়ে A-র যেকোনো <u>একটা</u> row-কে (বা যেকোনো <u>একটা</u> column-কে) গুণ করে দাও (বাকি সব অপরিবর্তিত রেখে), তবে determinant-টাও c দিয়ে গুণ হয়ে যাবে।

Example 11: শুরুতে ছিল ধরো $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. এর দ্বিতীয় row-টাকে 10 দিয়ে গুণ করলে হবে $B = \begin{bmatrix} 1 & 2 \\ 30 & 40 \end{bmatrix}$.

পরীক্ষা করে দেখো $|B| = 10|A|$ হচ্ছে কিনা।
SOLUTION: $|A| = 1 \times 4 - 2 \times 3 = -2$, আর $|B| = 1 \times 40 - 2 \times 30 = -20$. তার মানে সত্যিই $|B| = 10|A|$ হচ্ছে। ■

একইরকম ব্যাপার হবে কোনো column-কে গুণ করলেও। এতক্ষণ আমরা খালি <u>একটা</u> row বা column-কে গুণ করছিলাম। যদি দুটো row (বা column) নিয়ে কাজ করি, তবে গুণটাও দুবার হবে।

Exercise 16: যদি A-র প্রথম row-টাকে 5 দিয়ে আর দ্বিতীয় row-টাকে 3 দিয়ে গুণ করলে B পাওয়া যায়, আর $|A| = 10$ হয়, তবে $|B| = ?$ ■

ধরো একটা সংখ্যা c দিয়ে পুরো $A_{n \times n}$-কেই গুণ করলাম। তবে পাব $B = cA$. এখানে n-খানা row-ই c দিয়ে গুণ হয়েছে। তাই determinant-টাও n-বার c দিয়ে গুণ হবে, মানে $|B| = a^n|A|$.

Example 12: যদি $|A_{3\times3}| = 5$ হয়, তবে $|-A|$ কী হবে? যদি A-র সাইজ 4×4 হত, তবে?

SOLUTION: $|-A| = |(-1)A| = (-1)^3|A| = -5$.

যদি A-র সাইজ 4×4 হত, তবে $|-A| = (-1)^4|A| = 5$ হত। ■

Exercise 17: যদি $A_{n\times n}$ একটা skew symmetric matrix হয়, যেখানে n কোনো বিজোড় সংখ্যা, তবে দেখাও যে $|A| = 0$ হতে বাধ্য।
HINT: মনে রেখো skew symmetric মানে হল $A' = -A$. ■

35.4.3 Multiple যোগ করা

যদি একটা matrix-এর একটা row-এর multiple অন্য একটা row-এর সাথে যোগ করো, তবে determinant-টা বদলাবে না। একইরকম কথা column-দের বেলাতেও খাটে, সে তো বলাই বাহুল্য।

Example 13: ধরো $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ আছে। আমি এর প্রথম row-এর 10-গুণ দ্বিতীয় row-টার সাথে যোগ করে দিলাম। তবে পাব $B = \begin{bmatrix} 1 & 2 \\ 13 & 24 \end{bmatrix}$. পরীক্ষা করে দেখো $|A| = |B|$ হচ্ছে কিনা।
SOLUTION: $|A| = 1 \times 4 - 2 \times 3 = -2$, আর $|B| = 1 \times 24 - 2 \times 13 = -2$. ■

Exercise 18: আবার $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ নিচ্ছি। এবার প্রথম column-এর 5 গুণ দ্বিতীয় column-এ যোগ করে দাও। দেখো এতে determinant অপরিবর্তিত থাকছে কিনা। ■

DAY 36
প্রমাণ

গতকাল আমরা determinant-এর বিভিন্ন ধর্ম শিখেছি। আজকে তাদের প্রমাণগুলো শিখব। যারা বইটা প্রথমবার পড়ছ, তারা আজকের পড়াটায় খালি হালকা করে চোখ বুলিয়ে ছেড়ে দিতে পারো। তবে যারা অংকের যুক্তিতে মজা পাও, তাদের জন্য এটা আনন্দের দিন।

36.1 Matrix-এর পরিবর্তনের সঙ্গে determinant-এর পরিবর্তন

একটা matrix-এর পরিবর্তনের সঙ্গে তার determinant-এর পরিবর্তনের কয়েকটা সম্পর্ক শিখেছিলাম গতকাল। সেগুলোর প্রমাণ দেখা যাক।

36.1.1 Swap করা

কোনো দুটো row (বা কোনো দুটো column) যদি স্থানবিনিময় করে, তবে determinant-এর খালি চিহ্নটা উল্টে যায়। কেন এমনটা হয়, সেটা একটা ছোটো উদাহরণ নিয়ে পুরোটা লিখে দেখলেই বোঝা যাবে।

Example 14:

একটা 3×3 সাইজের matrix-এর জন্য পুরোটা লিখে ফেলি--

$$\begin{aligned} & \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ = \; & a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ & +a_{12}a_{23}a_{31} - a_{13}a_{21}a_{32} + \underline{a_{13}a_{22}a_{31}}. \end{aligned}$$

এবার প্রথম আর তৃতীয় row-এর স্থান বিনিময়ের পরে কী হচ্ছে, সেটা দেখালাম নীচে--

$$\begin{aligned} & \begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} \\ = \quad & \underline{a_{31}a_{22}a_{13}} + a_{31}a_{23}a_{12} + a_{32}a_{21}a_{13} \\ & -a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12} - a_{33}a_{22}a_{11}. \end{aligned}$$

ভালো করে খুঁটিয়ে লক্ষ্য কর যে, আগের ছয়টা জিনিসই ঘুরে ফিরে এসেছে, খালি প্রত্যেকটারই চিহ্ন গেছে উল্টে। যেমন ধরো, প্রথমটার বেলায় $a_{11}a_{22}a_{33}$-টা দ্বিতীয় ক্ষেত্রে হয়ে গেছে $a_{31}a_{22}a_{13}$ বোঝার সুবিধার জন্য নীচে দাগ দিয়ে দেখিয়েছি। একই term-টা কিন্তু swap-এর আগেও ছিল। সেখানে ওটাকে লিখেছিলাম $a_{13}a_{22}a_{31}$ হিসেবে। খালি সেখানে ওর চিহ্নটা উল্টো ছিল। ■

এবার বলি সর্বদাই এমনটা কেন হবে। তার জন্য আরেকটু বড় উদাহরণ নিয়ে ভাবলে সুবিধা। মনে করো একটা 8×8 সাইজের matrix-এর determinant-এর ফর্মুলাটা লিখছ। এতে 8!-খানা term থাকবে, তার মধ্যে একটা হল

$$a_{13}a_{22}a_{35}a_{48}a_{51}a_{64}a_{76}a_{87}.$$

এর চিহ্ন কী হবে? এখানে permutation-টা হল Fig 20-এর মত। এতে মোট ক্রসের সংখ্যা দেখা যাচ্ছে 9, একটা odd সংখ্যা। সুতরাং চিহ্নটা হবে '$-$'. এবার ধরো matrix-টার তৃতীয় এবং সপ্তম row-এর স্থানবিনিময় করা হল। তার ফলে যেখানে a_{3j} ছিল সেগুলো a_{7j}, আর a_{7j}-গুলো a_{3j} হয়ে যাবে। ফলে আমাদের term-টা হয়ে যাবে--

$$a_{13}a_{22}a_{75}a_{48}a_{51}a_{64}a_{36}a_{87}.$$

এর চিহ্ন সেই '$-$'-ই থাকবে। কিন্তু এই একই term তো শুরুর determinant-টাতেও ছিল। সেখানে এর চিহ্ন কী ছিল দেখা যাক। সেখানে ছবিটা ছিল Fig 21-এর মত। সবগুলো তীরই Fig 20-এর মতই আছে, খালি 3 আর 7 থেকে যে-দুটো তীর বেরিয়েছিল তাদের গোড়া দুটো স্থানবিনিময় করেছে। এদেরকে মোটা করে দেখিয়েছি। এখন মোট ক্রসের সংখ্যা 10, একটা even সংখ্যা। তার মানে এবার চিহ্নটা হবে '$+$'.

আমরা দ্বিতীয় ক্ষেত্রে নতুন করে ক্রসের সংখ্যা গুণে দেখলাম। কিন্তু সেটা না করেই বলে দেওয়া যেত যে, সংখ্যাটা একটা even সংখ্যা হবে। কীভাবে কাজটা করা যেত, তা বোঝার জন্য Fig 22 দেখো। এখানে বাঁদিকে দুটো তীর রয়েছে যারা পরস্পরকে ক্রস করে নি। এবার যদি মোটা তীরটার গোড়াটাকে সরিয়ে অন্য তীরটাকে টপকাই, তবে বুঝতেই পারছ যে, তীর দুটো এবার ক্রস করতে বাধ্য, যেমন দেখিয়েছি ডানদিকের ছবিটায়। আবার যদি গোড়াতেই তীর দুটো ক্রস করে থাকত (যেমন আছে Fig 23-এর বাঁদিকের ছবিটায়), তবে টপকানোর পর ক্রসটা চলে যেত। সুতরাং বুঝতেই পারছ যে, কোনো তীরের গোড়াকে আরেকটা তীরের গোড়ার উপর দিয়ে নিয়ে গেলে ক্রসের সংখ্যা সবসময়েই 1 বাড়ে বা 1 কমে।

এই কথাটা মাথায় রেখে এবার আবার Fig 20 এবং Fig 21-এর দিকে তাকাও। এখানে 3 নম্বর row থেকে বেরোনো তীরের গোড়াটাকে আমরা 7 নম্বর row-তে এনে ফেলেছি। মানে $4, 5$, আর 6-কে টপকেছি। মোট টপকানোর সংখ্যাকে যদি k বলি, তবে এখানে $k = 3$ হচ্ছে। এর ফলে ক্রসের সংখ্যা k-বার পরিবর্তিত হবে, প্রতি পরিবর্তনে 1 বাড়বে বা কমবে। একইভাবে

Fig 20

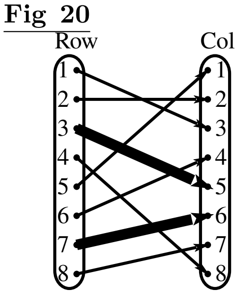


Fig 21

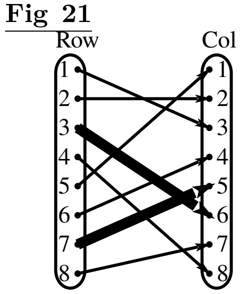


Fig 22

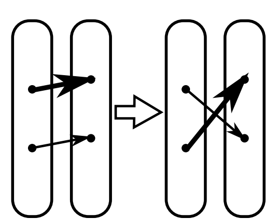
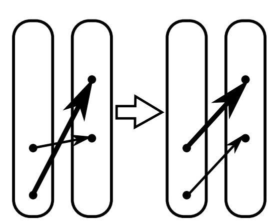


Fig 23



7 নম্বর row থেকে বেরোনো তীরের গোড়াটাকে নিয়ে ফেলেছি 3 নম্বরে, সুতরাং এখানেও ওই k-খানাকেই টপকাতে হল, অতএব ক্রসের সংখ্যা আরও k বার পরিবর্তিত হল। তার মানে মোট পরিবর্তন হল $2k$ পরিমাণ। এদিকে যে-দুটো তীরের গোড়ার স্থানবিনিময় করলাম, তাদের নিজেদের মধ্যে একটা ক্রসের সৃষ্টি হয়েছে (যদি গোড়ায় তাদের মধ্যে ক্রস থাকত, তবে স্থানবিনিময়ের ফলে সেটা চলে যেত)। সুতরাং আরও একটা পরিবর্তন পেলাম। ব্যস, আর কোথাও কোনো পরিবর্তন নেই, সুতরাং মোট পরিবর্তন হল $2k+1$ পরিমাণ, যেটা একটা odd সংখ্যা। অতএব গোড়ায় যদি even সংখ্যক ক্রস থেকে থাকে, তবে স্থানবিনিময়ের পরে সেই সংখ্যাটা odd হয়ে যাবে, আর গোড়ায় যদি odd নিয়ে শুরু করি, তবে পরে even হয়ে যাবে। অর্থাৎ মোদ্দা কথা দাঁড়ালো এই যে, swap-এর আগে ওই একই term-এর চিহ্নটা সবসময়েই উল্টো থাকতে বাধ্য, সে তুমি যা নিয়েই শুরু কর, আর যে দুটো row-কেই swap কর না কেন। প্রতিটা term-এর ক্ষেত্রেই যদি চিহ্নটা উল্টে যায়, তবে বুঝতেই পারছ যে, পুরো determinant-টার চিহ্নই উল্টে যাবে। এবং সেটাই তো দেখানোর কথা ছিল।

এই যুক্তি ব্যবহার করে একটা বিখ্যাত ধাঁধার সমাধান করা যায়। ধাঁধাটার একটা নাম হল 15 puzzle. ছোটোবেলায় আমরা অনেকেই এটা দেখেছি। বাচ্চাদের খেলনার দোকানে কিনতে পাওয়া যায়। সমাধানটার একটা আলোচনা পাবে পাশের QR code-টা scan করলেই। বা নীচের address-টা কোনো browser-এ লিখলেও হবে।
`https://arnabc74.github.io/linalg/15puzzle.html`

Exercise 19: ধরো একটা $n \times n$ সাইজের matrix আছে। যদি এর row-গুলোর ক্রমটা পুরো উল্টে দিই, তবে determinant-র চিহ্ন কীভাবে বদলাবে? উদাহরণস্বরূপ 3×3-এর বেলায় ক্রমটা হয়ে যাবে $3, 2, 1$ এবং 4×4-এর বেলায় $4, 3, 2, 1$.
HINT: ওল্টানোটাকে অনেকগুলো পরপর swap বলে ভাবো। ■

36.1.2 গুণ করা

এই জিনিসটা আমরা আগেই বলেছি--

> **THEOREM**
>
> Let A be any square matrix, and a be any scalar. If any row (or column) of A is multiplied by a, to produce a matrix B, then
> $$|B| = a|A|.$$

এটা কেন হয়, সেটা বোঝার জন্য আবার determinant-এর সংজ্ঞাটা মনে করা যাক। একটা $n \times n$ সাইজের matrix-এর determinant পাওয়া যায় $n!$-খানা term যোগ বিয়োগ করে। প্রত্যেকটা term তৈরী হয় matrix-টার n-খানা করে entry গুণ করে, এমনভাবে যাতে প্রতিটা term-এ প্রতিটা row এবং প্রতিটা column থেকে ঠিক একটা করেই entry থাকে।
এবার ধরো যেকোনো একটা row-এর প্রত্যেকটা entry-কে দ্বিগুণ করে দিলাম। এর ফলে determinant-টা কিভাবে বদলাবে? প্রত্যেকটা term-এই সেই row-এর ঠিক একজন করে প্রতিনিধি আছে। সে দ্বিগুণ হয়ে গেলে সেই term-টাও দ্বিগুণ হয়ে যাবে, কারণ একটা term-এর ভিতর সবাই গুণ হয়ে আছে। সবগুলো term-ই এভাবে দ্বিগুণ হয়ে যাওয়ার ফলে পুরো determinant-টাও দ্বিগুণ হয়ে যেতে বাধ্য! একই কথা খাটবে যেকোনো scalar দিয়ে গুণ করার ক্ষেত্রেই।

Exercise 20: যদি A একটা $n \times n$ সাইজের matrix হয়, তবে $|-A|$ কী $-|A|$ হবেই? ■

$$\begin{bmatrix} 1 & 3 & 4 & 0 \\ 9 & 3 & 5 & -1 \\ \mathbf{6} & \mathbf{4} & \mathbf{2} & \mathbf{1} \\ 4 & 9 & 3 & 10 \end{bmatrix}$$

Fig 24

$$\begin{bmatrix} 1 & 3 & 4 & 0 \\ 9 & 3 & 5 & -1 \\ \mathbf{2} & \mathbf{4} & \mathbf{1} & \mathbf{5} \\ 4 & 9 & 3 & 10 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 & 4 & 0 \\ 9 & 3 & 5 & -1 \\ \mathbf{4} & \mathbf{0} & \mathbf{1} & \mathbf{-4} \\ 4 & 9 & 3 & 10 \end{bmatrix}$$

Fig 25

Exercise 21: For a certain 9×9 matrix A, we have $|-A| = |A|$. Find $|A|$. ■

36.1.3 দুই টুকরো করা

এবার determinant-এর আরেকটা ধর্ম শিখব যা দিয়ে একটা determinant-কে ভেঙে দুই টুকরো করা যাবে। এর উল্লেখ আমরা গতকাল করিনি। একটা উদাহরণ নিয়ে শুরু করা যাক।

Example 15: ধরো Fig 24-এর matrix-টার যেকোনো একটা row নিলাম (মোটা করে দেখিয়েছি)। এই row-এর সব entry-কে ভেঙে দুটুকরো করলাম। যেমনভাবে ইচ্ছে করতে পারো, একটা উদাহরণ এইরকম--

$$(6, 4, 2, 1) = (2, 4, 1, 5) + (4, 0, 1, -4).$$

এবার Fig 25-এর মত দুটো matrix বানাও। বাকি row-গুলো দুটোতেই মূল matrix-টার মতই থাকবে, খালি ভাঙা row-টার বেলায় একটা matrix-এ থাকবে প্রথম টুকরোগুলো আর অন্যটায় যাবে দ্বিতীয় টুকরোগুলো। মজার কথা হল এই দুটো matrix-এর determinant যোগ করলে মূল matrix-টার determinant ফিরে পাবে! ■

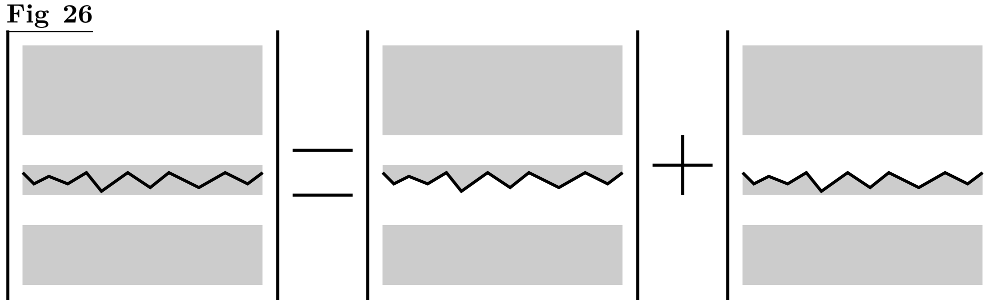
এটা প্রমাণ করা খুবই সহজ। কিন্তু প্রমাণের আগে Fig 26-এর মত ছবি দিয়ে মনে রাখলে সুবিধা হবে।
প্রমাণের জন্য গোড়ার matrix-টার নাম দিলাম ধরো A, যার (i, j)-th entry হল a_{ij}. এর মধ্যে k-th row-টাকে ভাঙা যাক। আমরা a_{kj}-কে ভেঙে যে-দুটো টুকরো পাব, তাদেরকে b_j আর c_j নাম দিই, মানে $a_{kj} = b_j + c_j$. এবার determinant-এর প্রতিটা term-এই এই ভাঙা row-টার প্রতিনিধি আছে। তার মানে term-গুলো দেখতে এইরকম--

$$a_{1f(1)} \cdots a_{kf(k)} \cdots a_{nf(n)}.$$

এখানে $a_{kf(k)}$ হল ভাঙা row-টার প্রতিনিধি। এবার ওটাকে ভেঙে লিখলে পাব--

$$a_{1f(1)} \cdots (b_{f(k)} + c_{f(k)}) \cdots a_{nf(n)}.$$

Fig 26



এবার ব্র্যাকেট তুললেই জিনিসটা দুটো term-এ ভাগ হয়ে যাচ্ছে--

$$a_{1f(1)} \cdots b_{f(k)} \cdots a_{nf(n)} + a_{1f(1)} \cdots c_{f(k)} \cdots a_{nf(n)}.$$

এইরকম সব term-কেই উপযুক্ত চিহ্ন লাগিয়ে যোগ করে দিলেই দুটো determinant পেয়ে যাবে--প্রথম ক্ষেত্রে ভাঙা row-টার জায়গায় b_j-গুলো থাকবে, আর দ্বিতীয় ক্ষেত্রে c_j-গুলো।

অবশ্যই একই যুক্তিতে তুমি দুভাগে না ভেঙে তিন কিংবা চার বা তারও বেশীসংখ্যক ভাগেও ভাঙতে পারো। Determinant-টাও ঠিক ততগুলো ভাগে ভেঙে যাবে। বলাই বাহুল্য যে, column-এর ক্ষেত্রেও একই ব্যাপার হবে।

Exercise 22: এমন তিনটে 2×2 সাইজের matrix দাও, A_1, A_2 আর A_3, যাতে এইটা হয়--

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ p & q & r \\ t & u & v \end{vmatrix} = \sum_1^3 a_i |A_i|.$$

■

36.1.4 Multiple যোগ করা

এই যে তিনটে জিনিস শিখলাম--determinant-কে ভেঙে দেওয়া, কোনো একটা row-কে একটা scalar দিয়ে গুণ করা, আর দুটো row সমান হয়ে গেলে determinant-এর শূন্য হয়ে যাওয়া--এদের মিলিয়ে একটা দরকারী জিনিস প্রমাণ হয়ে যায়--একটা row-এর কোনো multiple আরেকটা row-র সাথে গুণ করলে determinant-টা বদলায় না। Column-এর ক্ষেত্রেও একই ব্যাপার। প্রমাণটা খুবই সহজ। ধরো A একটা matrix যার row-গুলোকে নাম দিলাম $\vec{a}_1, ..., \vec{a}_n$. মানে

$$A = \begin{bmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_n' \end{bmatrix}.$$

আমরা $\vec{a}_i$-গুলোর transpose লিখেছি, কারণ আমরা শুয়ে থাকা row vector চাই। এবার যেকোনো দুটো row নাও, ধরো r-th এবং s-th row. আর যেকোনো একটা scalar নাও α. আমরা যদি $\alpha\vec{a}_r$ যোগ করে দিই $\vec{a}_s$-এর সাথে, তবে নতুন matrix-টা হবে

$$B = \begin{bmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_r' \\ \vdots \\ \vec{a}_s' + \alpha\vec{a}_r' \\ \vdots \\ \vec{a}_n' \end{bmatrix}.$$

এর determinant বার করব এইভাবে-- প্রথমে s-th row-টাকে দুইভাগে ভেঙে নেব, $\vec{a}_s'$ আর $\alpha\vec{a}_r'$. তাহলে determinant-

টাও দুইভাগে ভেঙে যাবে--

$$|B| = \begin{vmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_r' \\ \vdots \\ \vec{a}_s' \\ \vdots \\ \vec{a}_n' \end{vmatrix} + \begin{vmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_r' \\ \vdots \\ \alpha\vec{a}_r' \\ \vdots \\ \vec{a}_n' \end{vmatrix}.$$

এর মধ্যে প্রথমটা তো দেখাই যাচ্ছে $|A|$. দ্বিতীয়টা থেকে α-কে কমন নেওয়া যাবে--

$$|B| = |A| + \alpha \begin{vmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_r' \\ \vdots \\ \vec{a}_r' \\ \vdots \\ \vec{a}_n' \end{vmatrix}.$$

যে determinant-টা পড়ে রইল, তার r-th এবং s-th row সমান, সুতরাং ওই determinant-টা 0. অতএব $|B| = |A|$, ঠিক যা বলেছিলাম।

36.2 Elementary matrix-দের determinant

Elementary matrix-দের কথা মনে আছে কি? এরা হচ্ছে সেইসব matrix, যাদের দিয়ে বাঁদিক থেকে গুণ করে বিভিন্ন elementary row operation করা যায়। এরা তিন রকমের হয়--

$$S(i,j), \quad A(i,t,j) \quad \text{আর } M(i,t).$$

এদের determinant বার করা খুবই সহজ। মনে রেখো যে $S(i,j)$-কে পাওয়া যায় I-এর দুটো row-র স্থানবিনিময় করে। আমরা জানি $|I| = 1$, এবং swap করলে determinant-এর চিহ্নটা খালি উলটে যায়। সুতরাং

$$|S(i,j)| = -1, \quad (i \neq j).$$

একটা row-এর কোনো multiple আরেকটা row-তে যোগ করলে determinant বদলায় না। এদিকে $A(i,t,j)$-কে পাওয়া যায় I-এর i-th row-এর t-গুণ j-th row-এর সাথে যোগ করে। সুতরাং--

$$|A(i,t,j)| = |I| = 1, \quad (i \neq j).$$

কোনো row-কে কোনো সংখ্যা দিয়ে গুণ করে দিলে পুরো determinant-টাও সেই সংখ্যা দিয়ে গুণ হয়ে যায়। যেহেতু I-এর i-th row-কে t দিয়ে গুণ করে $M(i,t)$ পাওয়া গেছে, তাই--

$$|M(i,t)| = t.$$

36.3 Product

এবার আমরা দেখাব যে, A, B দুটো একই সাইজের square matrix হলে $|AB| = |A| \cdot |B|$ হয়। এর প্রমাণটা বেশ ঘোরালো, ধাপে ধাপে এগোব।

প্রথম ধাপে দেখাব যে, কোনো singular matrix-এর determinant শূন্য হতে বাধ্য। যুক্তিটা একটা উদাহরণ দিয়ে বলি। একটা square matrix-এর পক্ষে singular হওয়া মানে এর কোনো একটা row বাকী row-গুলোর linear combination. ধরো

চতুর্থ row $=2\times$ প্রথম row $-3\times$ দ্বিতীয় row.

এবার প্রথম row-এর 2-গুণ চতুর্থ row থেকে বিয়োগ করে দাও। এতে determinant বদলাবে না। এবার দ্বিতীয় row-এর 3-গুণ যোগ করে দাও চতুর্থ row-এর সাথে। এখনও determinant অপরিবর্তিত আছে। কিন্তু এখন তো চতুর্থ row-এর সব entry-ই শূন্য হয়ে গেছে। তাই পুরো determinant-ই শূন্য (যেহেতু প্রতিটা term-এই চতুর্থ row-এর একটা করে প্রতিনিধি আছে)। এদিকে determinant-টা বদলায় নি। তার মানে শুরুতেই determinant-টা শূন্যই ছিল।
এবার আমরা এটা ব্যবহার করে theorem-টার একটা বিশেষ ক্ষেত্র প্রমাণ করে ফেলতে পারব--যেখানে A আর B-এর মধ্যে অন্ততঃ একজনও singular. সেক্ষেত্রে $|A|\cdot|B|=0$. আবার AB-ও singular কারণ--

> A,B-র মধ্যে কেউ singular মানে তার rank হল $<n$ (যেখানে A আর B হল $n\times n$)। যেহেতু গুণ করলে rank বাড়তে পারে না, তাই $r(AB)<n$ হবেই।

সুতরাং $|AB|=0$. তাই এক্ষেত্রে $|AB|=|A|\cdot|B|$ হচ্ছে।
তাহলে বাকি রইল খালি সেই কেসটা যেখানে A,B দুজনেই nonsingular. এখানেও আমরা একটা বিশেষ ক্ষেত্র আগে প্রমাণ করে নেব, যেখানে A একটা elementary matrix. সেক্ষেত্রে প্রমাণটা খুব সহজ, কারণ elementary matrix আছেই মোটে তিন ধরণের--

- যদি $A=S(i,j)$ হয়, তবে AB মানে B-এর i-th row এবং j-th row দুটোর স্থানবিনিময় করে পাওয়া matrix. আমরা জানি যে, কোনো matrix-এর দুটো row-এর স্থানবিনিময় করলে determinant-র চিহ্ন উল্টে যায়। তাই $|AB|=-|B|$. এদিকে $|A|=|S(i,j)|=-1$. সুতরাং $|AB|=|A|\cdot|B|$ হচ্ছে।
- যদি $A=M(i,t)$ হয়, তবে AB মানে B-এর i-th row-কে t দিয়ে গুণ করে পাওয়া matrix. যেহেতু কোনো row-কে কোনো scalar দিয়ে গুণ করলে determinant-টাও সেই scalar দিয়ে গুণ হয়ে যায়, তাই $|AB|=t|B|$. এদিকে $|A|=|M(i,t)|=t$. সুতরাং এখানেও $|AB|=|A|\cdot|B|$ হচ্ছে।
- যদি $A=A(i,t,j)$ হয়, তবে AB মানে B-এর i-th row-এর t গুণ j-th row-এর সঙ্গে যোগ করে পাওয়া matrix. একটা matrix-এর একটা row-এর multiple আরেকটা row-তে যোগ করলে determinant বদলায় না। তাই $|AB|=|B|$. এদিকে $|A|=|A(i,t,j)|=1$. সুতরাং এখানেও $|AB|=|A|\cdot|B|$ হচ্ছে।

প্রমাণের অবশিষ্ট অংশের জন্য আমরা একটা অদ্ভুত কৌশল করব, RREF লাগাব। মনে আছে হয়তো যে, একটা nonsingular matrix-এর উপর Gauss-Jordan elimination লাগিয়ে গেলে শেষ পর্যন্ত I পাওয়া যায়, সেটাই হল RREF-টা। Gauss-Jordan elimination-এর প্রতিটা ধাপকে মনে করা যায় বাঁদিক থেকে একটা করে elementary matrix দিয়ে গুণ। যদি elementary matrix-গুলোকে $E_1,...,E_n$ বলি, তবে

$$E_n\cdots E_1A=I.$$

সুতরাং $A=E_1^{-1}\cdots E_n^{-1}$ হবে। অতএব A হল কিছু elementary matrix-এর গুণফল, কারণ elementary matrix-দের inverse-রাও ফের elementary matrix-ই হয়। একইভাবে B-ও কিছু elementary matrix-এর গুণফল। সুতরাং AB-ও তাই। এদিকে elementary matrix-দের গুণ করলে যে তাদের determinant-গুলোও গুণ হয়ে যায়, সে তো দেখলামই। তাই $|AB|=|A|\cdot|B|$ না হয়ে যায় কোথায়?

36.4 Block diagonal/triangular matrix

আমরা গতকাল শিখেছিলাম যে, block triangular matrix-দের ক্ষেত্রে diagonal block-গুলোর determinant গুণ করে দিলেই পুরো matrix-টার determinant পাওয়া যায়। যেহেতু block diagonal matrix-রাও block triangular matrix, তাই ওদের বেলাতেও একই নিয়ম খাটে। এই নিয়মের পিছনের যুক্তিটা এবার আমরা একটা উদাহরণ নিয়ে বুঝব।

Example 16: ধরো Fig 27-এর মত একটা block triangular matrix নিলাম। এখানে 'x'-গুলোর জায়গায় যা খুশি সংখ্যা থাকতে পারে (শূন্য হলেও আপত্তি নেই)। এটা 4×4 সাইজের, সুতরাং মোট $4!=24$-টা term থাকবে এর determinant-এ। এর মধ্যে অবশ্য অনেকেই শূন্য হতে বাধ্য। দেখা যাক কারা কারা nonzero হতে পারে।

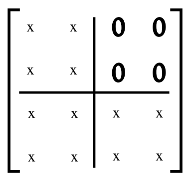


Fig 27

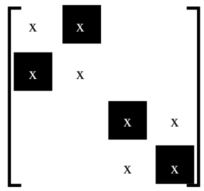


Fig 28

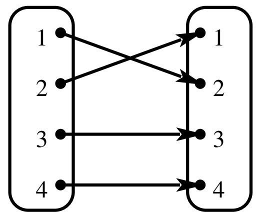


Fig 29

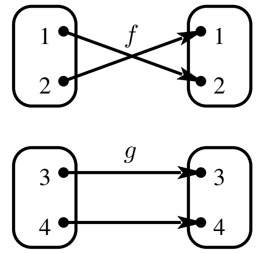


Fig 30

আমরা জানি যে, প্রত্যেকটা row থেকেই ঠিক একটা করেই entry আসবে প্রতিটা term-এ। প্রথম দুটো row থেকে দুটো entry নিতে হবে। বলাই বাহুল্য যে nonzero হবার জন্য দুজনকেই আসতে হবে প্রথম দুটো column থেকেই। এতেই প্রথম দুটো column খরচা হয়ে গেল। সুতরাং শেষের দুটো row-এর nonzero প্রতিনিধিরা আসতে বাধ্য শেষের দুটো column থেকে। সুতরাং term-গুলো হবে এইরকম--প্রথম diagonal block থেকে যতভাবে দুটো entry নেওয়া যায় (প্রতি row আর column থেকে ঠিক একটা করে নিয়ে), এবং এরকম যাবতীয় জিনিসের সঙ্গে গুণ হবে শেষের diagonal block-টা থেকে যতভাবে দুটো করে entry নেওয়া যায়। একটা এরকম term-এর সদস্যদের কালো করে দেখিয়েছি Fig 28-এ।
এবার চিহ্নের ব্যাপারটা ঠিক করা যাক। Fig 28-এর ক্ষেত্রে permutation-টা দেখিয়েছি Fig 29-এ। লক্ষ্য কর উপরের দুটো তীর এসেছে প্রথম diagonal block থেকে, আর নীচের দুটো তীর শেষের diagonal block-টা থেকে। এটা নিশ্চয়ই বুঝতে পারছ যে, কোনো nonzero term-এই উপরের দুটো তীর কখনই নীচের দুটোকে ক্রস করতে পারে না। সুতরাং মোট ক্রসিংএর সংখ্যা হবে উপরের দুটোর ক্রসিংএর সংখ্যা এবং নীচের দুটোর ক্রসিংএর সংখ্যার যোগফল, এক্ষেত্রে $1+0=1$. সুতরাং প্রত্যেকটা nonzero term-ই দেখতে হবে এইরকম--

$$sgn(f)a_{1f(1)}a_{2f(2)} \times sgn(g)a_{3g(3)}a_{4g(4)}.$$

এখানে f হল উপরের তীর দুটো দিয়ে তৈরী permutation, আর g হল নীচের তীর দুটো দিয়ে তৈরী। যেমন Fig 29-কে f, g-তে ভাঙলে পাব Fig 30. অতএব determinant-টা হবে

$$\sum_{f,g} sgn(f)a_{1f(1)}a_{2f(2)} \times sgn(g)a_{1g(1)}a_{2g(2)}.$$

এর মধ্যে f আর g-দের আলাদা করে করে নিলেই পেয়ে যাবে

$$\left[\sum_{f} sgn(f)a_{1f(1)}a_{2f(2)}\right] \times \left[\sum_{g} sgn(g)a_{1g(1)}a_{2g(2)}\right].$$

চৌকো ব্র্যাকেটের মধ্যেকার জিনিস দুটোই যে diagonal block দুটোর determinant, সেটা নিশ্চয়ই আর বলে দিতে হবে না? ■

36.5 Laplace expansion

আমরা গতকালই Laplace expansion-এর কথা শিখেছি। এবার দেখব সেটা কেন কাজ করে। Laplace expansion ব্যাপারটা ছিল এইরকম-- যদি $A_{n\times n}$-এর determinant বার করতে চাই, তবে সেটাকে লেখা যায়

$$|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} A_{ij}, \qquad (*)$$

যেখানে $i \in \{1, ..., n\}$ যেকোনো সংখ্যা, এবং A_{ij} হল A-র (i, j)-th minor. এটা ছিল i-th row বরাবর Laplace expansion. একইভাবে কোনো column বরাবরও Laplace expansion করা যায়।
আমরা দেখাব (*) কেন ঠিক। এর জন্য আমরা তিন ধাপে এগোব--

- প্রথম ধাপে দেখব যে, (*)-এর ডানদিকে A_{ij}-গুলোকে বড় করে লিখলে মোট $n!$-টাই term আসে,
- দ্বিতীয় ধাপে দেখব যে, প্রত্যেকটা term-এই প্রতিটা row এবং column থেকে ঠিক একজন করেই প্রতিনিধি আছে,
- তারপর দেখাব যে, চিহ্নগুলো সব determinant-এ যেমন যেমন থাকা উচিত, সেরকমই আছে।

প্রথম ধাপটা খুবই সোজা। ধরো $a_{ij}A_{ij}$-র কথাই নিলাম। এখানে A_{ij} হল একটা $n-1 \times n-1$ সাইজের matrix-এর determinant. সুতরাং এর মধ্যে $(n-1)!$-খানা term আছে। এইভাবে n-খানা $a_{ij}A_{ij}$-এর মত জিনিস থেকে মোট $n \times (n-1)! = n!$-খানা term পাচ্ছি, ঠিক যেমনটা পাওয়া দরকার ছিল। ভালো করে বুঝে দেখো যে সত্যিই $n!$-খানা আলাদা আলাদা term-ই পাচ্ছি, একই জিনিস দুই জায়গায় কখনো আসছে না, বা কোনো কিছু কাটাকাটি হয়ে যাচ্ছে না। যেমন ধরো $a_{i1}A_{i1}$-এর থেকে যে $(n-1)!$-খানা term আসছে তাদের বৈশিষ্ট্য হল যে এদের প্রত্যেকের মধ্যেই a_{i1} আছে, এবং a_{i2}, a_{i3} ইত্যাদিরা নেই (কারণ i-th row-এর বাকি সবাইকে কেটে দিয়েছি)। একই যুক্তিতে $a_{i2}A_{i2}$-তে যে $(n-1)!$-টা term আছে, তাদের মধ্যে a_{i2} আছে, এবং a_{i1}, a_{i3} ইত্যাদিরা নেই।

দ্বিতীয় ধাপ--এবার দেখাব যে, প্রত্যেকটা term-এই যে প্রতিটা row আর column-এর ঠিক একজন করেই প্রতিনিধি আছে। A_{i1} হল i-th row আর 1 নম্বর column বাদ দিলে যে matrix-টা পড়ে থাকে, তার determinant. Determinant-এর সংজ্ঞা অনুযায়ী এর মধ্যে যে $(n-1)!$-টা term আছে, তারা প্রত্যেকে i-th row বাদে প্রতিটা row এবং 1 নম্বর column বাদে প্রতিটা column থেকে ঠিক একটা করে প্রতিনিধি নিয়ে তৈরী। সুতরাং এদের সবার সঙ্গে যখন a_{i1} গুণ করছি, তখন যে $(n-1)!$-খানা term পাচ্ছি, তাদের মধ্যে i-th row আর 1 নম্বর column-এর প্রতিনিধিও এসে যাচ্ছে। একই রকম যুক্তি খাটছে অন্যান্য $a_{ij}A_{ij}$-দের জন্যও।

তৃতীয় ধাপে দেখতে হবে চিহ্নগুলো ঠিক থাকছে কী করে। এই জায়গাটা ধীরে ধীরে পড়। কঠিন নয়, কিন্তু প্রথমে মাথা একটু গুলিয়ে যাবার সম্ভাবনা। আবার $a_{ij}A_{ij}$ নিয়ে কাজ করি। আগেই বলেছি যে, এর মধ্যে $(n-1)!$-খানা term আছে, যাদের প্রত্যেকের মধ্যেই a_{ij} আছে। তীর চিহ্ন দিয়ে ভাবলে এর প্রত্যেকটা term তৈরী হয়েছে n-খানা তীর চিহ্ন দিয়ে, যার মধ্যে একটা তীর অবশ্যই i থেকে j-তে গেছে। Fig 31-এ একটা উদাহরণ দেখিয়েছি, যেখানে $i=4$ এবং $j=5$. ছবিতে i থেকে j-তে যাওয়া তীরটাকে মোটা করে দেখিয়েছি। এই তীরটার জন্ম হয়েছে a_{ij} থেকে, বাকি $n-1$-গুলোর জন্ম হয়েছে A_{ij} থেকে। এদিকে A_{ij} নিজেই একটা determinant, সুতরাং তার প্রতিটা term-এর একটা চিহ্ন আছে। সেই চিহ্নটা আসছে বাকি তীরগুলো নিজেদের মধ্যে কতবার ক্রস করেছে, তার উপর। প্রশ্ন হল-- মোটা তীরটা আসার ফলে মোট ক্রসের সংখ্যার কী পরিবর্তন হবে, অর্থাৎ ওই $n-1$-টা তীর মোটা তীরটাকে কতবার ক্রস করবে, even-বার নাকি odd-বার? এই প্রশ্নটার উত্তর বেশ মজার--

> যদি i, j দুজনেই even বা দুজনেই odd হয়, তবে even-বার ক্রস করবে, নইলে odd-বার।

লক্ষ্য কর উত্তরটা সম্পূর্ণই নির্ভর করছে i আর j-র উপরে, বাকি $n-1$-খানা তীর যা খুশি হতে পারে, তাতে কিছুই এসে যাচ্ছে না! এবার দেখি কেন এমন হয়।

Fig 31 দ্যাখো। মোটাটা বাদে বাকি $n-1$-খানাকে বলব রোগা তীর। রোগা তীরগুলোর নিজেদের মধ্যে ক্রসের সংখ্যা তো A_{ij}-র মধ্যেই ধরা আছে, আমাদের দেখতে হবে মোটা তীরটার জন্য কতগুলো নতুন ক্রসের জন্ম হয়। রোগা তীরগুলো মোটাটাকে দুভাবে ক্রস করতে পারে--হয় উপর থেকে নীচে, নয়তো নীচ থেকে উপরে। এবার ধাপে ধাপে এগোই--

- ধরো k-খানা তীর ক্রস করছে উপর থেকে নীচে। Fig 32-এ একটা উদাহরণ দেখিয়েছি, এখানে $i=4, j=5$ আর $k=2$. উপর থেকে নীচে ক্রস করা তীর দুটো ড্যাশ্ ড্যাশ্ দিয়ে এঁকেছি। এদের শেষ প্রান্ত আছে j-র নীচে।

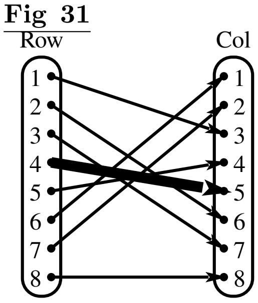
Fig 31


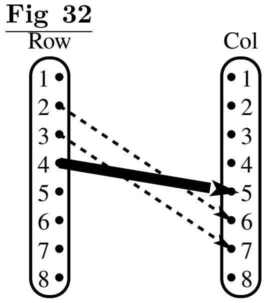
Fig 32


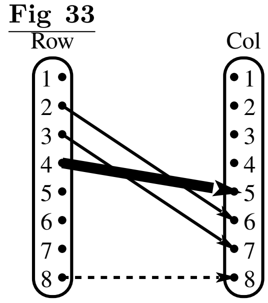
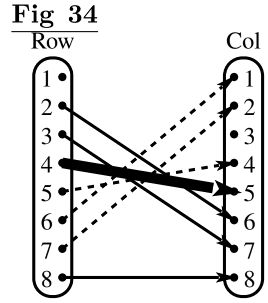
Fig 33


Fig 34


- এখন, j-র নীচে $n-j$-খানা জায়গা ছিল, তাদের মধ্যে k-খানা খরচা হয়ে গেল, পড়ে রইল $n-j-k$-খানা। আবার Fig 32-এর সঙ্গে মিলিয়ে নাও। $j=5$-এর নীচে $n-j=3$-টে জায়গা ছিল, $k=2$-খানা খরচা হয়ে পড়ে রইল 1-টা।
- এই সব জায়গায় যে তীর গুলো এসেছে তারা শুরু হয়েছে i-এর নীচ থেকে (নইলে তারাও মোটা তীরটাকে ক্রস করত উপর থেকে নীচে)। Fig 33-এর উদাহরণে এরকম একটাই তীর আছে, ড্যাশ্ ড্যাশ্ দিয়ে দেখিয়েছি।
- এবার দ্যাখো যে, i-এর নীচে $n-i$-খানা জায়গা ছিল, তার থেকে $n-j-k$-খানা খরচা হয়ে হয়ে গেল। পড়ে রইল $(n-i)-(n-j-k)=k-i+j$-খানা। এদের থেকে যে সব তীর বেরোবে তাদের বাধ্য হয়েই j-র উপরে যেতে হবে, মানে মোটা তীরটাকে নীচ থেকে উপরে ক্রস করতে হবে। Fig 34-এ এরকম তীরের সংখ্যা হল 3, যাদের ড্যাশ্ ড্যাশ্ দিয়ে এঁকেছি।
- সুতরাং মোটা তীরটাকে মোট ক্রস করা হচ্ছে $k+(k-i+j)=2k-i+j$ বার।
- যেহেতু $2k$ সব সময়েই even, তাই এই সংখ্যাটা even নাকি odd হবে, সেটা নির্ভর করছে $i-j$-র even/odd হওয়ার[2] উপরে। যদি i, j দুজনেই even হয় বা দুজনেই odd হয়, তবে এটা even হবে, নইলে odd.

বেশ সুন্দর যুক্তিপরম্পরা, তাই না?

DAY 37
Laplace expansion-এর প্রয়োগ

এবার আমরা Laplace expansion-এর দুটো প্রয়োগ শিখব।

37.1 Finding inverse

ধরো A হল একটা 3×3 সাইজের matrix–

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}.$$

তাহলে আমরা Laplace expansion থেকে শিখেছি যে

$$a_{11}C_{11}+a_{12}C_{12}+a_{13}C_{13}=|A|,$$

যেখানে C_{ij}-রা হল cofactor (মানে চিহ্নগুলো ওদের মধ্যে ধরা আছে)। এটা যে কোনো 3×3 সাইজের matrix-এর ক্ষেত্রেই প্রযোজ্য। সুতরাং যদি প্রথম row-এর জায়গায় $(4,3,7)$ নিই তবে হবে

$$\begin{vmatrix} 4 & 3 & 7\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}=4C_{11}+3C_{12}+7C_{13}.$$

একইভাবে যদি প্রথম row-এর জায়গায় দ্বিতীয় row-টার একটা কপি তুলে বসিয়ে দিই, তবে হবে

$$\begin{vmatrix} a_{21} & a_{22} & a_{23}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}=a_{21}C_{11}+a_{22}C_{12}+a_{23}C_{13}.$$

[2]আমরা আগে $(i+j)$-র even/odd হওয়ার কথা বলেছিলাম। সেটা একই ব্যাপার, কারণ $(i-j)$ আর $(i+j)$ দুজনে একই সঙ্গে even বা odd হয়।

কিন্তু একটু ভালো করে দেখলেই বুঝবে যে, বাঁদিকের determinant-টা আসলে 0, কারণ এখানে দ্বিতীয় row আর প্রথম row সমান। সুতরাং

$$a_{21}C_{11} + a_{22}C_{12} + a_{23}C_{13} = 0.$$

একই যুক্তিতে যদি প্রথম row-এর জায়গায় তৃতীয় row-টা তুলে বসিয়ে দিতাম, তবে পেতাম

$$a_{31}C_{11} + a_{32}C_{12} + a_{33}C_{13} = 0.$$

এখানে আমরা 3×3 সাইজের matrix নিয়ে কাজ করছিলাম। কিন্তু যেকোনো $n \times n$ সাইজের matrix-এর ক্ষেত্রেই একই ব্যাপার--

- Laplace expansion থেকে পাই $\sum_{k=1}^{n} a_{ik}C_{ik} = |A|$,
- আর যদি $i \neq j$ নিই, তবে $\sum_{k=1}^{n} a_{ik}C_{jk} = 0$. কারণ এটা হল A-র j-th row-টাকে মুছে তার জায়গায় i-th row-টার একটা কপি বসিয়ে দিলে যে matrix-টা পাওয়া যায়, তার determinant.

হাতেকলমে একটা উদাহরণ কষে দেখে নিই।

Example 17: ধরো একটা 2×2 সাইজের matrix নিলাম--

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

তুমি অন্য যা খুশি matrix নিয়েও দেখতে পারো। এখানে $|A| = -2$, এবং

$$a_{11} = 1, \quad a_{12} = 2, \quad a_{21} = 3, \quad a_{22} = 4,$$

আর cofactor-গুলো হল

$$C_{11} = 4, \quad C_{12} = -3, \quad C_{21} = -2, \quad C_{22} = 1.$$

মনে রেখো এগুলো হল cofactor, অর্থাৎ চিহ্নটা এদের মধ্যে ঢোকানোই আছে।
এবার $i, j = 1, 2$ বসিয়ে $\sum_{k=1}^{2} a_{ik}C_{jk}$ বার করে দ্যাখো তো কী হয়।
Solution:

- যদি $i = j = 1$ নিই, তবে

$$a_{11}C_{11} + a_{12}C_{12} = 1 \times 4 + 2 \times (-3) = -2 = |A|.$$

- যদি $i = j = 2$ নিই, তবে

$$a_{21}C_{21} + a_{22}C_{22} = 3 \times (-2) + 4 \times 1 = -2 = |A|.$$

- যদি $i = 1, \quad j = 2$ নিই, তবে

$$a_{11}C_{21} + a_{12}C_{22} = 1 \times (-2) + 2 \times 1 = 0.$$

- যদি $i = 2, \quad j = 1$ নিই, তবে

$$a_{21}C_{11} + a_{22}C_{12} = 3 \times 4 + 4 \times (-3) = -2 = 0.$$

ঠিক এমনটাই হওয়ার কথা ছিল। ■

এবার আমরা এই জিনিসটাকেই কায়দা করে সাজিয়ে লিখব। তার জন্য একটা জিনিস একটু মনে করে নিই, যেটা একেবারে প্রথম অধ্যায়ে শিখেছিলাম--

> যদি $C_{m\times n}$ আর $B_{n\times p}$ দুটো matrix হয়, তবে তাদের গুণফল AB-র (i,j)-th entry-টা হবে $\sum_{k=1}^n a_{ik}b_{kj}$.

এইটা ব্যবহার করে চট্ করে একটা অংক করে নিই--

Example 18: ধরো A, B দুজনেই $n\times n$ সাইজের। তবে AB'-এর (i,j)-th entry কী হবে?

SOLUTION: যদি B'-এর (i,j)-th entry-কে b'_{ij} লিখি তবে বুঝতেই পারছ যে $b'_{ij} = b_{ji}$.
সুতরাং AB'-এর (i,j)-th entry হবে

$$\sum_{k=1}^n a_{ik}b'_{kj} = \sum_{k=1}^n a_{ik}b_{jk}.$$

■

এবার এই অংকে b_{jk}-র জায়গায় C_{jk} কল্পনা কর। তাহলে AB'-এর (i,j)-th entry হবে $\sum_{k=1}^n a_{ik}C_{jk}$, এবং আমরা একটু আগেই দেখিয়েছি যে

$$\sum_{k=1}^n a_{ik}C_{jk} = \begin{cases} |A| & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}.$$

এখানে যাকে B' বলেছি তার পোশাকি নাম হল **adjugate** of A বা সংক্ষেপে $\text{adj}(A)$.

DEFINITION: Adjugate

For any square matrix $A_{n\times n}$, its **adjugate** $\text{adj}(A)$, is defined as the $n\times n$ matrix with (i,j)-th entry equal to the (j,i)-th cofactor of A.

আগেকার দিনে লোকে এটাকে adjoint of A বলত, কিন্তু এখন[3] adjoint বলতে অন্য জিনিস[4] বোঝায়। এখানে একটা ব্যাপারে সতর্ক। Adjugate-এর (i,j)-th entry-টা কিন্তু (j,i)-th cofactor বললাম, i আর j-র ক্রমটা উল্টে গেছে! যদি গুলিয়ে যায়, তবে একটা সহজ উদাহরণ দেখে নিই।

Example 19: $\text{adj}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = ?$

SOLUTION: এখানে $C_{11}=4$, $C_{12}=-2$, $C_{21}=-3$ এবং $C_{22}=1$. তাই

$$\text{adj}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}.$$

■

[3] পুরোনো অর্থে adjoint বোঝাতে চাইলে লোকে অনেক সময়ে classical adjoint বলে।
[4] সেটা আমরা এই বইতে পরে শিখব।

Exercise 23: $\text{adj}\begin{bmatrix} 5 & -2 \\ 0 & 1 \end{bmatrix} = ?$ ■

আমরা এক্ষুণি যা আলোচনা করলাম, তা থেকে পাচ্ছি যে, $A \cdot \text{adj}(A)$ হল এমন একটা matrix, যার যাবতীয় diagonal entry-ই হল $|A|$. আর বাকি সব entry হল 0. তার মানে এইরকম দেখতে--

$$A \cdot \text{adj}(A) = \begin{bmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |A| \end{bmatrix} = |A| \cdot I_n.$$

একটা theorem-এর আকারে লিখে রাখা যাক--

> **THEOREM**
>
> For any square matrix A we have
>
> $$A \cdot \text{adj}(A) = |A| \cdot I.$$

মাথা গুলিয়ে যাচ্ছে? তবে আবার আমাদের উদাহরণটা দেখে নিই--

Example 20: আমাদের উদাহরণে আমরা $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ নিয়ে কাজ করছিলাম। এখানে $|A| = -2$, আর

$$C_{11} = 4, \quad C_{12} = -3, \quad C_{21} = -2, \quad C_{22} = 1.$$

এদেরকে সাজিয়ে প্রথমে একটা matrix বানাই--

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}.$$

এবার এর transpose নিলেই পাব

$$\text{adj}(A) = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}.$$

এবার $A \cdot \text{adj}(A)$ বার করে দেখো কী পাও।

SOLUTION: এখানে

$$A \cdot \text{adj}(A) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = -2 \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = |A| \cdot I.$$

■

এই theorem থেকে কী করে A^{-1} বার একটা পথ বেরোয় দেখি। যদি $|A| \neq 0$ হয়, তবে লিখতে পারি

$$A \cdot \frac{1}{|A|}\text{adj}(A) = I,$$

অর্থাৎ $\frac{1}{|A|}\text{adj}(A)$ এমন একটা matrix যা দিয়ে A-কে গুণ করলে I হচ্ছে। তার মানেই তো হল A-টা nonsingular, এবং এ থেকে A^{-1} বার করার একটা কায়দাও পাওয়া যাচ্ছে--

THEOREM

For any nonsingular matrix A we have

$$A^{-1} = \frac{1}{|A|}\operatorname{adj}(A).$$

অবশ্য একটু বড় matrix-এর বেলায় এই কায়দাটা মোটেই সহজ কিছু নয়। কারণ এই কায়দায় inverse বার করতে গেলে এক রাশ determinant বার করতে হয়। নীচের $|A|$-টা তো আছেই, তার উপর আবার $\operatorname{adj}(A)$-র প্রতিটা entry-ই তো আসলে একটা করে determinant.
যাই হোক, একটা ছোটো উদাহরণ করে দেখাই।

Example 21: যদি $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ হয়, যেখানে $ad - bc \neq 0$, তবে A^{-1} কী হবে?

SOLUTION: এখানে $\operatorname{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. সুতরাং

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

যেটা আমাদের পূর্বপরিচিত। ■

Adjugate-এর ধারণা কেমন তোমার মাথায় ঢুকেছে, সেটা বোঝা যাবে নীচের অংক দুটো করতে পারলে।

Exercise 24: দেখাও যে $|A| = 0$ হলে $|\operatorname{adj}(A)| = 0$ হতে বাধ্য।
HINT: Proof by contradiction লাগাও। যদি $|\operatorname{adj}(A)| \neq 0$ হয়, তবে $\operatorname{adj}(A)$-এর inverse আছে। এবার $A \cdot \operatorname{adj}(A) = |A| \cdot I$ থেকে চিন্তা করো। ■

Exercise 25: দেখাও যে, যদি A-র সাইজ $n \times n$ হয়, তবে $|\operatorname{adj}(A)| = |A|^{n-1}$.
HINT: দুটো কেসে ভেঙে এগোও--যখন $|A| \neq 0$ আর যখন $|A| = 0$. আগের অংকটা কাজে লাগতে পারে। ■

37.2 Cramer's rule

যদি একটা system of linear equations দেওয়া থাকে $A\vec{x} = \vec{b}$, যেখানে $A_{n\times n}$ হল nonsingular, তবে এই system-টার unique solution থাকবে। সেটা বার করার একটা অদ্ভুত কায়দা আছে, তার নাম **Cramer's rule**. অদ্ভুত এই কারণে যে, চট্ করে দেখলে বোঝা কঠিন কায়দাটা কেন কাজ করে। কায়দাটা খুব যে সুবিধাজনক কিছু এমন নয়, এর জন্য $n+1$-খানা determinant বার করতে হয়, প্রত্যেকটাই $n \times n$ সাইজের matrix-এর। কায়দাটা কেন কাজ করে, সেটা আমরা পরে প্রমাণ করব। আপাততঃ জিনিসটা কী বলি।

Cramer's rule

Let $A_{n\times n} = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$ be any nonsingular matrix. Then the unique solution to the system

$$A\vec{x} = \vec{b}$$

is given by $\vec{x} = (x_1, ..., x_n)'$, where

$$x_i = \frac{\begin{vmatrix} \vec{a}_1 & \cdots & \vec{a}_{i-1} & \vec{b} & \vec{a}_{i+1} & \cdots & \vec{a}_n \end{vmatrix}}{|A|}.$$

ব্যাপারটা কী হচ্ছে বুঝে নেওয়া যাক। প্রত্যেকটা x_i হবে একটা determinant-কে আরেকটা determinant দিয়ে ভাগ করে তৈরী। নীচের determinant-টা সব সময়েই $|A|$. উপরের determinant-টায় একটু কারিকুরি আছে। যদি তুমি x_i বার করতে চাও, তবে A-র i-th column-টার জায়গায় $\vec{b}$ বসিয়ে দাও। এইভাবে যে নতুন matrix-টা পাবে, সেটার determinant-ই রয়েছে উপরে। নীচের অংকটায় একটা উদাহরণ দেখালাম।

Example 22: Solve by Cramer's rule

$$\begin{aligned} x + y + z &= 1 \\ ax + by + cz &= k \\ a^2x + b^2y + c^2z &= k^2, \end{aligned}$$

where $a \neq b \neq c$.

SOLUTION:

The given system is

$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ k \\ k^2 \end{bmatrix}.$$

এখানে Van der Monde matrix-এর ছড়াছড়ি। এদের determinant বার করার কায়দা তো জানোই--

$$\begin{vmatrix} 1 & 1 & 1 \\ k & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (a-b)(a-c)(b-c),$$

$$\begin{vmatrix} 1 & 1 & 1 \\ k & b & c \\ k^2 & b^2 & c^2 \end{vmatrix} = (k-b)(k-c)(b-c),$$

$$\begin{vmatrix} 1 & 1 & 1 \\ a & k & c \\ a^2 & k^2 & c^2 \end{vmatrix} = (a-k)(a-c)(k-c),$$

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & k \\ a^2 & b^2 & k^2 \end{vmatrix} = (a-b)(a-k)(b-k).$$

$\therefore$ By Cramer's rule the solution is

$$x = \frac{\begin{vmatrix} 1 & 1 & 1 \\ k & b & c \\ k^2 & b^2 & c^2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix}} = \frac{(k-b)(k-c)(b-c)}{(a-b)(a-c)(b-c)} = \frac{(k-b)(k-c)}{(a-b)(a-c)},$$

and

$$y = \frac{\begin{vmatrix} 1 & 1 & 1 \\ a & k & c \\ a^2 & k^2 & c^2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix}} = \frac{(a-k)(a-c)(k-c)}{(a-b)(a-c)(b-c)} = \frac{(a-k)(k-c)}{(a-b)(b-c)},$$

and

$$z = \frac{\begin{vmatrix} 1 & 1 & 1 \\ a & b & k \\ a^2 & b^2 & k^2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix}} = \frac{(a-b)(a-k)(b-k)}{(a-b)(a-c)(b-c)} = \frac{(a-k)(b-k)}{(a-c)(b-c)}.$$

■

Exercise 26: Cramer's rule দিয়ে solve করো--

$$x + 2y = 5, \quad 3x + 4y = 11.$$

■

37.2.1 Cramer's rule-এর প্রমাণ

আমাদের system of equations-টা হল--

$$A\vec{x} = \vec{b},$$

যেখানে A হল $n \times n$ সাইজের। যেহেতু A একটা nonsingular matrix বলা আছে, তাই এই system-টার একটা unique solution থাকবে--

$$\vec{x} = A^{-1}\vec{b}.$$

এক্ষুণি শিখলাম যে $A^{-1} = \frac{1}{|A|}\text{adj}(A)$ লেখা যায়। তার মানে

$$\vec{x} = \frac{1}{|A|}\text{adj}(A)\vec{b}.$$

এবার আমরা $\text{adj}(A)\vec{b}$ অংশটাকে একটু খুঁটিয়ে দেখব। তার জন্য একটা 3×3 উদাহরণ নিয়ে চিন্তা করলে সুবিধা হবে। যদি A-র সাইজ 3×3 হয়, তবে

$$\text{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix}.$$

লক্ষ্য কর এর (i,j)-th entry কিন্তু C_{ji} (মোটেই C_{ij} নয়!)। যদি $\vec{b}$-র entry-গুলোকে b_i বলি, আর $\vec{x}$-এর entry-গুলোকে বলি x_i, তবে

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

এবার entry-গুলোকে একে একে দেখি। প্রথমে দেখি x_1 কত হয়--

$$x_1 = \frac{1}{|A|}\left[b_1C_{11} + b_2C_{21} + b_3C_{31}\right].$$

একটু চিন্তা করে দেখলেই বুঝবে যে, চৌকো ব্র্যাকেটের মধ্যে যেটা আছে সেটা আসলে এই determinant-টা--

$$\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}.$$

যদি পরিষ্কার না হয়, তবে প্রথম column বরাবর Laplace expansion করে দ্যাখো। একইভাবে

$$x_2 = \frac{1}{|A|}\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix},$$

আর

$$x_3 = \frac{1}{|A|}\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

যদি A-র সাইজ $n \times n$ হত, তাহলেও একই জিনিস খাটবে। ধরো A-র column-গুলোকে $\vec{a}_j$ নাম দিলাম, মানে

$$A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}.$$

তবে A-র i-th column-র জায়গায় $\vec{b}$ বসিয়ে দিলে পাব

$$\begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_{i-1} & \vec{b} & \vec{a}_{i+1} & \cdots & \vec{a}_n \end{bmatrix}.$$

এটার determinant-কে $|A|$ দিয়ে ভাগ করলেই পেয়ে যাব x_i, মানে

$$x_i = \frac{1}{|A|}\begin{vmatrix} \vec{a}_1 & \cdots & \vec{a}_{i-1} & \vec{b} & a_{i+1} & \cdots & \vec{a}_n \end{vmatrix}.$$

এবং এটাই ছিল Cramer's rule.

Answers

1. দুইক্ষেত্রেই '$-$'। **2.** $+1$. **3.** যদি $n = 2k$ বা $2k+1$ হয়, তবে $sgn(f) = (-1)^k$. কারণ এ পাশের k-খানা তীর ওপাশে, আর ও পাশের k-খানা তীর এ পাশে আসবে। যদি n বিজোড় হয়, তবে মাঝের তীরটা যথাস্থানে থাকবে। **4.** হ্যাঁ। **5.** 0. **6.** 6. **7.** Upper triangular-দের জন্য প্রমাণের কায়দাটা সংক্ষেপে বলছি। Lower triangular-দের জন্যও মূল যুক্তিটা একই থাকবে। যদি matrix টা $n \times n$ হয়, তবে n-এর উপরে induction লাগাও। $n = 1$ হলে কিছুই করার নেই, কারণ 1×1 সাইজের matrix মানে একটাই সংখ্যা, এবং সেটাই তার নিজের determinant. এবার ধরে নাও যে, যেকোনো $(n-1) \times (n-1)$ সাইজের matrix-এর জন্য প্রমাণ হয়ে গেছে। তবে $n \times n$ সাইজের matrix-এর determinant-টাকে প্রথম row দিয়ে Laplace expansion করো। যেহেতু matrix-টা lower triangular, তাই প্রথম row-তে খালি একটা entry-ই nonzero. সুতরাং determinant-টা হবে সেইটা গুণ তার cofactor. কিন্তু cofactor-টা তো একটা $(n-1) \times (n-1)$ সাইজের matrix-এর lower triangular matrix. তাই induction hypothesis লাগাতে পারবে। **8.**

> Let the given matrix be A.
>
> Let A_{ij} be its (i,j)-th cofactor. By Laplace's method
>
> $$|A| = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} + a_{14}A_{14}.$$
>
> Now
>
> $$A_{11} = \begin{vmatrix} a & d & -c \\ -d & a & b \\ c & -b & a \end{vmatrix} = \cdots = a(a^2+b^2+c^2+d^2),$$

এখানে ওই $\cdots$ মানে কিছু মামুলী ধরণের ধাপ বাদ দিয়ে গিয়েছি।
একইভাবে--

> $$A_{12} = -\begin{vmatrix} -b & d & -c \\ -c & a & b \\ -d & -b & a \end{vmatrix} = \cdots = b(a^2+b^2+c^2+d^2),$$
>
> $$A_{13} = \begin{vmatrix} -b & a & -c \\ -c & -d & b \\ -d & c & a \end{vmatrix} = \cdots = c(a^2+b^2+c^2+d^2),$$
>
> $$A_{14} = -\begin{vmatrix} -b & a & d \\ -c & -d & a \\ -d & c & -b \end{vmatrix} = \cdots = d(a^2+b^2+c^2+d^2).$$

So

$$\begin{aligned}|A| &= a^2(a^2+b^2+c^2+d^2)+b^2(a^2+b^2+c^2+d^2)+\\ &\quad c^2(a^2+b^2+c^2+d^2)+d^2(a^2+b^2+c^2+d^2)\\ &= (a^2+b^2+c^2+d^2)^2,\end{aligned}$$

as required.

9.

By Laplace's formula applied to the first row we have

$$\begin{aligned}D_4 &= \begin{vmatrix} 1 & x & 0 & 0\\ x & 1 & x & 0\\ 0 & x & 1 & x\\ 0 & 0 & x & 1\end{vmatrix}\\ &= 1\times\begin{vmatrix} 1 & x & 0\\ x & 1 & x\\ 0 & x & 1\end{vmatrix} - x\begin{vmatrix} x & x & 0\\ 0 & 1 & x\\ 0 & x & 1\end{vmatrix}\end{aligned}$$

এবার লক্ষ্য কর যে, প্রথম determinant-টা আসলে D_3. আবার দ্বিতীয় determinant-টাকে যদি প্রথম row বরাবর Laplace expansion করি, তবে পাব

$$x\begin{vmatrix} 1 & x\\ x & 1\end{vmatrix}.$$

তার মানে সব মিলিয়ে D_4 দাঁড়ালো গিয়ে--

$$\begin{aligned}&= 1\times D_3 - x^2\begin{vmatrix} 1 & x\\ x & 1\end{vmatrix}\\ &= 1\times D_3 - x^2 D_2.\end{aligned}$$

Hence

$$D_4 - D_3 + x^2 D_2 = 0,$$

as required.

যদি $x=\frac{1}{2}$ নিই, তবেই পরের অংশের matrix-টা পাব।

Taking $x=\frac{1}{2}$, we have

$$D_4 - D_3 + \tfrac{1}{4}D_2 = 0.$$

Now

$$D_2 = \begin{vmatrix} 1 & x\\ x & 1\end{vmatrix} = 1 - x^2 = \frac{3}{4}.$$

Also,

$$D_3 = \begin{vmatrix} 1 & x & 0 \\ x & 1 & x \\ 0 & x & 1 \end{vmatrix} = 1 - 2x^2 = \frac{1}{2}.$$

So

$$D_4 = D_3 - \tfrac{1}{4}D_2 = \frac{1}{2} - \frac{1}{4} \times \frac{3}{4} = \frac{5}{16}.$$

10. প্রথম row বরাবর Laplace expansion করেও করতে পারো, বা সরাসরি determinant-এর সংজ্ঞা ব্যবহার করেও করতে পারো। যদি $n = 2k$ বা $2k+1$ হয়, তবে $D_n = (-1)^k a_1 \cdots a_n$. **11.** $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ আর $B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ নিতে পারো। **12.** Rank হল 2, কারণ প্রথম আর তৃতীয় row-এর average (গড়) নিলে দ্বিতীয় row-টা হয়। কিন্তু প্রথম আর তৃতীয় row পরস্পরের multiple নয়।
Rank-এর সাইজের একটা submatrix হল $\begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}$, যার determinant হচ্ছে $-4 \neq 0$. **13.** এখানে দুবার স্থানবিনিময় হয়েছে বলে ভাবা যায়, শুরুতে প্রথম row দুটো, তারপর শেষের row দুটো। তাই চিহ্নটা দুবার উল্টে ফের সোজা হয়ে যাবে। তাই $|B| = a$. **14.** ধরো matrix-টা হল A. ওই row দুটোর স্থানবিনিময়ের পর হল B. তবে $|B| = -|A|$. এদিকে row দুটো তো একই, তাই ওদের স্থানবিনিময়ের ফলে আসলে matrix-টা বদলায় নি। মানে $B = A$. সুতরাং $|A| = -|A|$, অর্থাৎ $|A| = 0$. **15.** না, যেমন $\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} \neq 0$ যদিও এর প্রথম row-টাকে খাড়া করে দিলেই প্রথম column-টা পাওয়া যায়। **16.** $|B| = 10 \times 3 \times 5 = 150$. **17.** $A' = -A$, তাই $|A'| = |-A|$ অর্থাৎ $|A| = (-1)^n|A| = -|A|$, যেহেতু n হল বিজোড়। সুতরাং $|A| = 0$. **18.** নতুন matrix-টা হবে $\begin{bmatrix} 1 & 7 \\ 3 & 19 \end{bmatrix}$. এর determinant হল $1 \times 19 - 7 \times 3 = -2$. তার মানে determinant-টা অপরিবর্তিত থাকছে। **19.** যদি $n = 2k$ বা $2k+1$ হয়, determinant-টার চিহ্ন ওল্টাবে যখন k বিজোড় হবে। যদি k জোড় হয়, তবে determinant-টা অপরিবর্তিত থাকবে। **20.** যদি n বিজোড় হয়, তবেই। **21.** $|A| = 0$. **22.** একটা উত্তর এরকম, $A_1 = \begin{bmatrix} q & r \\ u & v \end{bmatrix}$, $A_2 = \begin{bmatrix} r & p \\ v & t \end{bmatrix}$ এবং $A_3 = \begin{bmatrix} p & q \\ t & u \end{bmatrix}$. **23.** $\begin{bmatrix} 1 & 0 \\ 2 & 5 \end{bmatrix}$. **24.** $A \cdot \mathrm{adj}(A) = |A| \cdot I$ থেকে পাচ্ছি $A \cdot \mathrm{adj}(A) = O$. যেহেতু $\mathrm{adj}(A)$-কে nonsingular ধরেছি, তাই এর মানে হল $A = O$. কিন্তু সেক্ষেত্রে $\mathrm{adj}(A) = O$ হয়ে যাচ্ছে। অথচ আমরা ধরে নিয়েছিলাম যে $|\mathrm{adj}(A)| \neq 0$. তাই contradiction. **25.** $A \cdot \mathrm{adj}(A) = |A| \cdot I$ থেকে দুপাশের determinant নিয়ে $|A| \cdot |\mathrm{adj}(A)|^n = |A|^n$. যদি $|A| \neq 0$ হয়, দুপাশ থেকে $|A|$ কেটে নিলেই হবে। আর $|A| = 0$ হলে আগের অংকের যুক্তি লাগানো যাবে। **26.** $x = \dfrac{\begin{vmatrix} 5 & 2 \\ 11 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}} = \frac{-2}{-2} = 1$ এবং

$$y = \frac{\begin{vmatrix} 1 & 5 \\ 3 & 11 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}} = \frac{-4}{-2} = 2.$$

Chapter VI
Quadratic and bilinear forms

DAY 54
গোড়ার কথা

54.1 Quadratic forms

54.2 কী জিনিস

আমরা determinant-এর ইতিহাস আলোচনার সময়ে বলেছিলাম কীভাবে determinant-এর ধারণাটা একরকম "পণ্ডিতের খেলনা" হয়েই বহুদিন ছিল। সেই অবস্থার অবসান ঘটালেন Gauss, যিনি determinant দিয়ে প্রথম এমন একটা কাজ করে দেখালেন, যেটা determinant ছাড়া সহজে করা যাচ্ছিল না। সে আমলে অনেক গণিতজ্ঞই নিছক গণিতজ্ঞ ছিলেন না, তাঁরা বিজ্ঞানের অন্যান্য শাখাতেও কাজ করতেন। যেমন Gauss নানা গুরুত্বপূর্ণ কাজ করেছেন physics এবং astronomy-তে। Astronomy-তে Gauss-এর অন্যতম অবদান ছিল বিভিন্ন গ্রহ ইত্যাদির orbit (কক্ষপথ) বার করা। তোমরা হয়তো জানো যে, এই orbit-গুলো হয় ellipse, parabola ইত্যাদি জাতীয়। এই প্রসঙ্গে বলে রাখি, এই ধরণের curve-গুলোকে বলে একেকটা **conic section**. সৌর জগতের (বা তার বাইরেও) যে সব জিনিস একই পথে বার বার পাক খায়, তারা সাধারণতঃ ellipse-এর মত চলে, আর যারা একবার দেখা দিয়েই দূর শূন্যে মিলিয়ে যায়, তাদের চলার পথটা হয় parabola-র মত। এখন মনে করো দূরবীন দিয়ে আকাশে একটা নতুন কিছু জিনিসকে ছুটে আসতে দেখা গেল। যতদিন অব্ধি দেখা গেল, তার পথটা পণ্ডিতরা খুব ভালো করে দূরবীন দিয়ে মেপে রাখলেন। এবার প্রশ্ন হল এটা কী দেখলাম--নতুন একটা গ্রহ জাতীয় কিছু, যেটা ফিরে ফিরে আসবে, নাকি এমন কিছু যেটা আর কখনোই আসবে না? এমন হতেই পারে যে, এটা হয়তো তিনশ' বছর পরে পরে ফিরে আসে। তিনশ' বছর আগে তো আর দূরবীন ছিল না, তাই মানুষ এর কথা আগে জানতেও পারে নি। কিন্তু যদি তিনশ' বছর পরে ফিরে আসে, তবে তো সে কথা লিখে রাখতে হবে, যাতে আগামী দিনের মানুষ হিসেব মিলিয়ে দেখতে পারে। এদিকে আমাদের হাতে তো আছে খালি জিনিসটার orbit-এর অল্প খানিকটা অংশ, যেটুকু ওই কয়দিনে দূরবীন দিয়ে মাপা গেছে। সুতরাং Gauss-এর চিন্তা ছিল কী করে একটা orbit-এর equation থেকে বোঝা যায় সেটা কোন ধরণের conic section. এই অব্ধি পড়ে তুমি হয়তো ভাবছ--এটা আর শক্ত কী, ellipse-এর equation তো হয় এইরকম--

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

আর parabola হয়

$$y = 4ax^2,$$

আর hyperbola হয়

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

এদের তো চোখে দেখেই আলাদা করা যায়! তবে Gauss এর মধ্যে এত মাথা ঘামানোর কী পেলেন? আসলে এখানে ব্যাপারটা আরেকটু শক্ত। এই যে তিনটে equation-এর চেহারা তুমি দিলে, এরা কিছু বিশেষ ধরণের ellipse, parabola

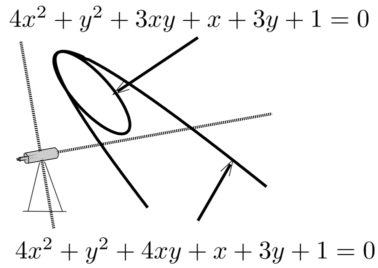


Fig 1

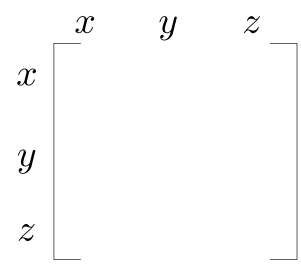


Fig 2

আর hyperbola-র ক্ষেত্রেই খাটে--যখন এদের axis-গুলো x-axis বা y-axis বরাবর থাকে। কিন্তু যদি কোনো ellipse, parabola বা hyperbola তেরছা হয়ে থাকে, তবে কিন্তু তার equation বেশ বিচ্ছিরি দেখতে হতে পারে। এখন বুঝতেই পারছ যে, আমরা যখন দূরবীন দিয়ে কোনো orbit মাপছি, তখন আমাদের x-axis আর y-axis নিচ্ছি দূরবীনের পক্ষে সুবিধাজনক অবস্থানে। কিন্তু আকাশের গ্রহ, ধূমকেতু ইত্যাদির তো আর আমাদের দূরবীনের axis-এর সঙ্গে তাল রেখে চলার কোনো দায় নেই! সুতরাং এক্ষেত্রে তেরছা[1] conic section-দের নিয়ে কাজ করা ছাড়া গত্যন্তর নেই (Fig 1)। তাহলে প্রথম কাজ হল যেকোনো conic section-এর equation-এর একটা general form বার করা, যেটা সবক্ষেত্রেই কাজ করবে, parabola, ellipse, circle, hyperbola তেরছা হোক বা না হোক। এরকম একটা equation হল--

$$ax^2 + by^2 + cxy + dx + ey + f = 0.$$

খুব একটা জটিল কিছু বলা যায় না, কিন্তু এতে দুটো সমস্যা--

- এক, যেকোনো parabola,ellipse ইত্যাদির equation-ই এই চেহারার হলেও আরও কিছু ভুলভাল জিনিসও এর মধ্যে ঢুকে পড়েছে, যেমন ধরো $a = b = f = 1$ আর $c = d = e = 0$ নিলে হয়

  $$x^2 + y^2 + 1 = 0.$$

  বলাই বাহুল্য যে, এরকম কোনো $x, y \in \mathbb{R}$ সম্ভব নয়, কারণ $x^2, y^2 \geq 0$ হওয়ায় বাঁদিকটা ≥ 1 হবে, এদিকে ডানদিকে তো 0 হয়ে বসে আছে!

- আর দ্বিতীয় সমস্যা হল $a, b, ..., f$-এর কোনো value বলে দিলে চট্ করে বোঝা কঠিন জিনিসটা ellipse নাকি parabola বা অন্য কিছু।

এই দুটো সমস্যার সমাধান করার কাজেই অপ্রত্যাশিতভাবে determinant কাজে এসে গেল। কীভাবে বলি। Gauss প্রথমে এই equation-টা matrix দিয়ে লিখলেন এইভাবে--

$$\begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} a & c/2 & d/2 \\ c/2 & b & e/2 \\ d/2 & e/2 & f \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0.$$

গুণটা করেই দেখো! এইবার তিনি দেখলেন যে, মাঝের matrix-টার determinant-টা যদি 0 হয়, তবে জিনিসটা circle, ellipse, parabola, hyperbola কিছুই হয় না। যদি determinant-টা $\neq 0$ হয়, তবে জিনিসটা কী হবে, সেটা নির্ভর করে খালি এই matrix-টার উপরে--

$$\begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix}.$$

[1]অংকের ভাষায় এদের বলে oblique conic section.

- এর determinant যদি < 0 হয়, তবে আমরা পাব একটা hyperbola,
- যদি $= 0$ হয়, তবে parabola,
- আর > 0 হলে ellipse বা circle হতে পারে। কিন্তু এখানেও আবার একটা ভুলভাল কেস ঢুকে পড়ে। সেটা হল এই যে, $a+b$-এর চিহ্ন যদি প্রথম determinant-টার উল্টো হয়, তবে সেরকম কোনো $x, y \in \mathbb{R}$ পাওয়া অসম্ভব। আর যদি $a = b$ এবং $c = 0$ হয়, তবে পাব একটা circle.

Gauss এতসব কী করে বার করলেন, সে আলোচনা করার জায়গা এটা নয়। এখানে বলার কথা হল, এই প্রথম একটা দরকারী কাজে determinant লাগল, যেটা অন্যভাবে করার কোনোই সহজ কায়দা চোখে পড়ছিল না। এর সাহায্যে একটা conic section-এর প্রকৃতি determine (নির্ধারণ) করা গেল বলে Gauss-ই প্রথম একে determinant নাম দেন। সেই নামই আজও চলে আসছে।

এই কাজটা করতে গিয়ে Gauss একটা নতুন জিনিস বানিয়ে ফেলেছিলেন, যেটা অংকের দুনিয়ায় conic section-এর প্রকৃতি নির্ধারণ করার চেয়ে অনেক বেশী গুরুত্বপূর্ণ বলে স্বীকৃত। সেটা হল **quadratic form.** লক্ষ্য কর, Gauss কীভাবে conic section-এর equation-টাকে matrix দিয়ে লিখেছিলেন। এখানে যেন তিনটে প্রধান জিনিস ছিল x, y আর 1. এদের square-গুলোর coefficient ছিল a, b আর f. তাদেরকে লিখেছিলাম diagonal বরাবর। আর বাকিরা ছিল এদের বিভিন্ন গুণফল। এদেরকে সমান দুইভাগ করে দুইদিকে লিখেছিলাম। মোটামুটিভাবে একেই বলে quadratic form, খালি এক্ষেত্রে আমরা $x, y, 1$-এর বদলে x, y, z নিয়ে কাজ করি। অবশ্য তিনটে variable-ই নিতে হবে এমন কোনো কথা নেই, যেকোনো n-সংখ্যক নেওয়া যায়।

প্রথমে $n = 3$ নিয়ে একটা উদাহরণ দেখি।

Example 1: ধরো এই জিনিসটাকে matrix দিয়ে লিখতে চাই--

$$3x^2 + 8y^2 - 24z^2 + 10yz + 10zx - 14xy.$$

SOLUTION: কী করে কাজটা করা যাবে, সেটা Fig 2 থেকে Fig 5 পর্যন্ত দেখানো হয়েছে। এখানে তিনটে variable x, y, z আছে, তাই একটা 3×3 matrix নিয়ে শুরু করেছি (Fig 2)। গোড়াতে matrix-টা ফাঁকা, খালি row আর column-গুলো x, y, z দিয়ে লেবেল করেছি। এবার square term-গুলোর coefficient-দের বসাব diagonal বরাবর (Fig 3)। তারপর প্রত্যেকটা product term-এর coefficient-কে ভেঙে সমান দুই ভাগ করেছি, যেমন $-14xy$-এর -14-কে অর্ধেক করে পেয়েছি দুটো -7. এই দুটোকে matrix-এর দু জায়গায় বসিয়েছি (Fig 4)। একই ভাবে বাকী দুটো product term বসিয়ে পেয়েছি পুরো matrix-টা (Fig 5)।

এটা তিনটে জিনিসের গুণফল, প্রথমে একটা শোয়ানো vector, তারপর একটা square matrix এবং তারপরে ফের প্রথম vector-টাকেই খাড়া করে লেখা। এই তিনটেকে গুণ করে দেখো, ঠিক $3x^2 + 8y^2 - 24z^2 + 10yz + 10zx - 14xy$ পাবে। ■

Fig 3

$3x^2 + 8y^2 - 24z^2$

$$\begin{array}{c|ccc|} & x & y & z \\ \hline x & 3 & & \\ y & & 8 & \\ z & & & -24 \end{array}$$

Fig 4

$-14xy$

$$\begin{array}{c|ccc|} & x & y & z \\ \hline x & 3 & -7 & \\ y & -7 & 8 & \\ z & & & -24 \end{array}$$

Fig 5

$10yz + 10zx$

$$\begin{array}{c|ccc|} & x & y & z \\ \hline x & 3 & -7 & 5 \\ y & -7 & 8 & 5 \\ z & 5 & 5 & -24 \end{array}$$

ভালো করে বুঝে নাও যে, এই কায়দাটা সব সময়েই কাজ করবে, এবং এর ফলে যে matrix-টা পাবে, সেটা একটা symmetric matrix হবে। যদি n-খানা variable থাকে, $x_1, ..., x_n$, তবে তাদের নিয়ে quadratic form মানে হল যতরকম $x_i x_j$ জাতীয় জিনিস হয়, তাদের linear combination. যদি $i = j$ হয়, তবে x_i^2 জাতীয় জিনিস পাবে। Matrix-টা বানানো শুরু করবে diagonal entry-গুলো থেকে। ওর i-th diagonal entry হবে x_i^2-এর coefficient. আর যদি $i \neq j$ নাও, তবে a_{ij} আর a_{ji} মিলে $x_i x_j$-র coefficient হবে। আবার symmetry রক্ষার জন্য $a_{ij} = a_{ji}$-ও চাই। তাই a_{ij} আর a_{ji} দুজনকেই নেব $x_i x_j$-র coefficient-এর অর্ধেক করে।
এই কথাটা গুছিয়ে লিখলেই quadratic form-এর সংজ্ঞাটা পেয়ে যাবে--

DEFINITION: **Quadratic form**

By a **quadratic form** in the variables, $x_1, ..., x_n$, we mean a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(x_1, ..., x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} x_i x_j,$$

where r_{ij}'s are real numbers.
The quadratic form can be written as $\vec{x}' A \vec{x}$, where

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

and $A_{n\times n}$ has (i, j)-th entry

$$a_{ij} = \begin{cases} r_{ii} & \text{if } i = j \\ \frac{r_{ij}}{2} & \text{if } i \neq j \end{cases}$$

Example 2: এদের মধ্যে কারা quadratic form আর কারা নয়?

(1) $(x + y)(3y - 2x)$ **(2)** $(x + 1)^2$ **(3)** $x^2 y$ **(4)** $(x_1 + 2x_2 + \cdots + nx_n)^2$ **(5)** $\frac{1}{x^2+y^2}$

1. হ্যাঁ, quadratic form, কারণ ব্র্যাকেট তুলে গুণ করলে প্রতিটা term-এই ঠিক দুটো variable-এর গুণফল থাকবে।
2. না, $(x + 1)^2 = x^2 + 2x + 1$-এর মধ্যে শেষের দুটো term মোটেই quadratic নয়।
3. না, $x^2 y$-এর মধ্যে <u>তিনটে</u> variable গুণ হয়ে আছে, দুটো x আর একটা y.
4. এটা একটা quadratic form, কারণ এখানেও ব্র্যাকেট তুলে গুণ করলে প্রতিটা term-এই ঠিক দুটো variable-এর গুণফল থাকবে।
5. না, $x^2 + y^2$ অবশ্যই একটা quadratic form, কিন্তু ওটা এখানে denominator-এ চলে যাওয়ায় পুরো জিনিসটা আর quadratic form হচ্ছে না!

■

Exercise 1: এদের মধ্যে কারা quadratic form আর কারা নয়?
(1) $x^2-y^2+(x+z)(3x+2y-z)$ **(2)** $(x+y+1)^2$ **(3)** $(xy)^2$ **(4)** $(x_1+x_2\cdots+x_n)^2$ **(5)** $x^2+\frac{1}{x^2}$
■

Exercise 2: নীচের quadratic form-গুলোকে matrix দিয়ে লেখো।

1. $f(x,y)=x^2-3y^2+4xy$.
2. $f(x,y)=(x+3y)^2$.
3. $f(x,y,z)=xy+yz+zx$.
4. $f(x,y,z)=(x+3y)^2$. এখানে ফর্মুলাটায় কোনো z নেই। কিন্তু তাও যেহেতু $f(x,y,z)$ লেখা হয়েছে, matrix-টা 3×3 সাইজের হবে।

■

যেকোনো quadratic form-কেই matrix-এর আকারে লেখা যায়। কিন্তু একই quadratic form-কে কি একাধিকভাবে matrix দিয়ে লেখা যায়? যদি matrix-টাকে symmetric নাও, তবে উত্তর হল--না, প্রত্যেকটা quadratic form-এর জন্য ঠিক একটাই symmetric matrix পাওয়া যাবে। সেটা একটা theorem-এর আকারে লিখে রাখা যাক।

THEOREM

If $A_{n\times n}$ and $B_{n\times n}$ are two symmetric matrices corresponding to the same quadratic form, *i.e.*,

$$\forall \vec{x}\in\mathbb{R}^n \quad \vec{x}'A\vec{x}=\vec{x}'B\vec{x},$$

then we must have

$$A=B.$$

Proof:

Let $\vec{x}=(x_1,...,x_n)'$ and $A\equiv((a_{ij}))$ and $B\equiv((b_{ij}))$.
Then, by the given condition,

$$\forall x_1,...,x_n\in\mathbb{R} \quad \sum_i\sum_j a_{ij}x_ix_j=\sum_i\sum_j b_{ij}x_ix_j.$$

Hence, equating coefficients of x_i^2, we have $a_{ii}=b_{ii}$.
Also, by equating coefficients of x_ix_j for $i\neq j$, we have $a_{ij}+a_{ji}=b_{ij}+b_{ji}$.
But, since A and B are symmetric, hence $a_{ij}=a_{ji}$ and $b_{ij}=b_{ji}$.
So we have $2a_{ij}=2b_{ij}$, and hence $a_{ij}=b_{ij}$.
Thus, $A=B$, as required.

∎

[Q.E.D]

এই কারণে যখনই আমরা একটা quadratic form-কে matrix দিয়ে লিখি, তখন symmetric matrix-ই ব্যবহার করি।

Example 3: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). Let A be a matrix of the quadratic form $(x_1 + 2x_2 + \cdots + nx_n)^2$. Then the sum of the entries of A is

(A) $\sum n$ (B) $\sum n^2$ (C) $\sum n^3$ (D) $\frac{n(n+1)}{2}$

(2021.CC12.1i)

SOLUTION: এই অংকে "<u>a</u> matrix of the quadratic form" বলেছে। তার মানে symmetry শর্তটা চাপাচ্ছে না। সেটা এই অংকে সত্যিই দরকারী নয়। Symmetric হোক আর না হোক উত্তরটা একই হবে। তাই মনে মনে symmetric ভেবে নিলেও আপত্তি নেই।

প্রথম option তিনটে যেভাবে দিয়েছে তার কোনো মানে হয় না। সম্ভবতঃ প্রশ্নকর্তার মাথায় এইরকম কিছু একটা ঘুরছিল যে, $\sum n$ হল $\frac{n(n+1)}{2}$. অবশ্যই এর মানে হয় না, কারণ বাঁদিকে n রয়েছে summation-এর ভিতরে, মানে sum-টা n-এর উপরে হচ্ছে, সেক্ষেত্রে উত্তরটা আবার n-এর উপরে নির্ভর করে কী করে? আসলে বলা উচিত ছিল $\sum_{i=1}^n k$. সেই হিসেবে আমরা option-গুলোকে ধরে নিচ্ছি--

(A) $\sum_1^n k$ (B) $\sum_1^n k^2$ (C) $\sum_1^n k^3$ (D) $\frac{n(n+1)}{2}$

এই অংকটা নানাভাবে করা যায়। একেবারে বুদ্ধি না খাটিয়ে ফাঁকি মেরে করার কায়দা হল ছোটো ছোটো n বসিয়ে দেখা। যেমন ধরো, $n = 2$ বসালেই quadratic form-টা হয়ে যাবে $(x_1 + 2x_2)^2 = x_1^2 + 4x_2^2 + 4x_1x_2$. এর (symmetric) matrix-টা হল $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. এর সবগুলো entry যোগ করলে হচ্ছে $1 + 2 + 2 + 4 = 9$. পরপর option-গুলো পরীক্ষা করলেই দেখবে খালি (C)-ই পড়ে থাকছে।

তবে এটা নেহাত ফাঁকি হল। যদি একটা option থাকত "None of the above", অমনি এই কায়দায় কাজ হত না। কিন্তু যদি একটু বুদ্ধি খাটাই, তবে আরেকটা কায়দা বেরোয়, যেটাতে পরিশ্রম ফাঁকির কায়দাটার চেয়েও কম! মনে রেখো একটা quadratic form আসলে $\vec{x}'A\vec{x} = \sum_i \sum_j a_{ij}x_ix_j$. এখান থেকে বার করতে বলেছে $\sum_i \sum_j a_{ij}$. তার মানে $x_1 = \cdots x_n = 1$ বসিয়ে দিলেই হল।

> Let $f(x_1, ..., x_n) = \vec{x}'A\vec{x} = (x_1 + 2x_2 + \cdots + nx_n)^2 = \sum_i \sum_j a_{ij}x_ix_j$.
>
> Then $\sum_i \sum_j a_{ij} = f(1, ..., 1) = (1 + 2 + \cdots + n)^2 = \left(\frac{n(n+1)}{2}\right)^2$.
>
> So the correct option is (C).

এবার আরো একটা কায়দা বলি। এই কায়দায় পরিশ্রম একটু বেশী হলেও বাড়তি সুবিধা হল A-টা দেখতে কীরকম সেটা জানা যাবে (যদিও এই অংকে A-টা মোটেই বার করতে বলেনি)। আমরা A-কে symmetric ধরে নিচ্ছি। এখানে

$$x_1 + 2x_2 + \cdots + nx_n = \underbrace{\begin{bmatrix} 1 & 2 & \cdots & n \end{bmatrix}}_{\vec{b}'} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

এবার একটা ছোট্টো ম্যাজিক করব। লক্ষ্য করো এটাকে আমরা উল্টো দিক দিয়েও লিখতে পারি, $\vec{x}'\vec{b}$. এবার যখন square করব, তখন এই দুটো রূপকে গুণ করব, মানে

$$\vec{x}'A\vec{x} = (x_1 + 2x_2 + \cdots + nx_n)^2 = \vec{x}'\vec{b}\vec{b}'\vec{x}.$$

যেহেতু A-কে symmetric ধরেছি, তাই--

$$A = \vec{b}\vec{b}' = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix} \begin{bmatrix} 1 & 2 & \cdots & n \end{bmatrix},$$

যার (i, j)-th entry হল $a_{ij} = ij$.

তাই $\sum_i \sum_j a_{ij} = \sum_i \sum_j ij = \left(\sum_1^n i\right)\left(\sum_1^n j\right) = \left(\frac{n(n+1)}{2}\right)^2$. ■

নীচের অংকটাও একই যুক্তি দিয়ে করা যাবে।

Exercise 3: একটা quadratic form দিচ্ছি $(\sum_1^9 i^2 x_i)^2$. এটার matrix-টাকে যদি $A_{9\times 9}$ বলি, তবে এর প্রথম চারটে row আর প্রথম চারটে column দিয়ে তৈরী submatrix-টার সব entry-র যোগফল কত হবে? ■

ধরো একটা $f : \mathbb{R}^n \to \mathbb{R}$ আছে, যেটা একটা quadratic form. যদি কোনো $\vec{x}$ নাও, এবং সেটাকে a-গুণ করো $(a \int \mathbb{R})$, তবে $f(\vec{x})$-টা a^2-গুণ হয়ে যাবে। এর উল্টো দিকটাও কি সত্যিই? সেটা নিয়েই নীচের অংকটা।

Exercise 4: ধরো একটা $f : \mathbb{R}^n \to \mathbb{R}$ এমন একটা function যাতে

$$\forall \vec{x} \in \mathbb{R}^n \quad \forall a \in \mathbb{R} \quad f(a\vec{x}) = a^2 f(\vec{x}).$$

তবে কি f একটা quadratic form হতে বাধ্য? ■

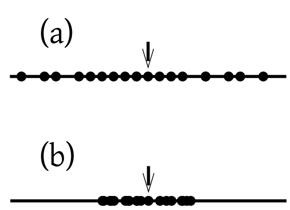
Exercise 5: দেখাও যে, $\mathbb{R}^n$-এর উপরে defined যাবতীয় quadratic form-রা মিলে একটা vector space তৈরী হয়। এর dimension কত?

HINT:

একটা quadratic form-কে একটা symmetric matrix বলে ভাবতে পারো। সুতরাং প্রশ্নটা হল এটা দেখানো যে, যাবতীয় $n \times n$ সাইজের symmetric matrix-রা মিলে একটা vector হয়, এবং তার dimension কী। ■

DAY 55
কোথায় লাগে?

Fig 6



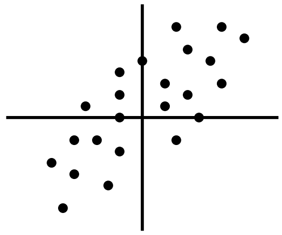


Fig 7

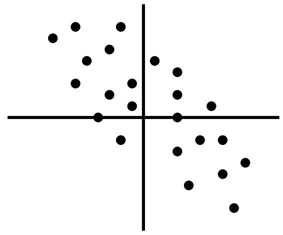


Fig 8

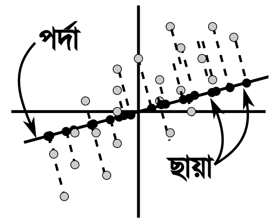


Fig 9

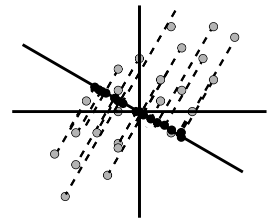


Fig 10

Quadratic form-এর ধারণাটা বিজ্ঞানের বিভিন্ন শাখায় কাজে লাগে। আমরা এরকম তিনটে প্রয়োগ আলোচনা করব। একটা statistics-এ, একটা physics-এ, এবং multivariate calculus-এ।

55.1 Statistics

ধরো কিছু সংখ্যা আছে $x_1, ..., x_n$. আমরা এখানে real number নিয়ে কাজ করব। যদি $\mathbb{R}$-কে একটা লম্বা লাইন বলে ভাবো, তবে এই সংখ্যাগুলো হবে সেই লাইনের উপর n-খানা বিন্দু। দুটো উদাহরণ দিয়েছি Fig 6-এ। যদি জিজ্ঞাসা করি এই দুটো ছবির মধ্যে কোনটাতে point-গুলো বেশী ছড়িয়ে ছিটিয়ে আছে, তবে তোমার উত্তর হবে--প্রথম ছবিটাতে। এটা চোখে দেখেই অনায়াসে বলা যায়। কিন্তু ধরো আমরা এই ছড়িয়ে ছিটিয়ে থাকার পরিমাণটাকে অংক কষে মাপতে চাই। তার একটা সহজ পথ এইরকম--প্রথমে সংখ্যাগুলোর average (গড়) বার কর, মানে

$$\overline{x} = \frac{1}{n}\sum_1^n x_i.$$

এইটা তোমায় বলবে যে point-গুলোর ভীড়ের কেন্দ্রটা কোথায়। ছবিতে তীর চিহ্ন দিয়ে দেখিয়েছি এদের। এবার দেখো point-গুলো এই কেন্দ্র থেকে কতটা দূরে আছে। সেটা মাপার একটা কায়দা হল $(x_i - \overline{x})$. এটা অবশ্য positive বা negative দুইই হতে পারে। কিন্তু চিহ্ন নিয়ে আমাদের মাথাব্যথা নেই, আমরা খালি দূরত্বটুকু জানলেই খুশি। তাই চিহ্নটাকে তাড়ানোর জন্য আমরা ওটাকে square করে নিয়ে পাব $(x_i - \overline{x})^2$. এবার আমরা এই (squared) দূরত্বগুলোর average নেব, যাতে মোটামুটিভাবে point-গুলো কেন্দ্র থেকে কত দূরে আছে, তার একটা আন্দাজ পাই--

$$\frac{1}{n}\sum_1^n (x_i - \overline{x})^2.$$

Statistics-এ এ জিনিসটার একটা নাম আছে, একে বলে **variance,** এর সাহায্যে মাপা যায় x_i-রা কতটা ছড়িয়ে-ছিটিয়ে আছে, যেমন Fig 6 (a)-এর বেলায় variance হবে Fig 6 (b)-এর variance-এর চেয়ে বেশী। এই উদাহরণটা ছিল $\mathbb{R}$-এ, মানে dimension যেখানে 1. এবার ধরো point-গুলো আছে $\mathbb{R}^2$-তে। Fig 7 দেখো। এক্ষেত্রে কী করে অংক কষে "ছড়িয়ে ছিটিয়ে থাকা" মাপবে? এখানে প্রশ্ন খালি "কতটা" ছড়িয়ে ছিটিয়ে আছে, সেটাই না। "কীভাবে" ছড়িয়ে ছিটিয়ে আছে, সেটাও একটা বড় প্রশ্ন। যেমন Fig 7 আর Fig 8-এ ছড়িয়ে ছিটিয়ে থাকার পরিমাণ একইরকম, কিন্তু বিন্যাসের দিকটা অন্য। ছড়িয়ে-ছিটিয়ে থাকার পরিমাণ এবং দিক দুটোকেই একইসঙ্গে প্রকাশ করতে quadratic form কাজে লাগে। প্রথমে ছবি দিয়ে বোঝাচ্ছি। ধরো যেকোনো একটা দিক নিলাম। অংকে একটা দিক বোঝানো হয় সেটা বরাবর একটা unit vector $\vec{w}$ দিয়ে। এই দিক বরাবর point-গুলো কতটা ছড়িয়ে-ছিটিয়ে আছে, সেটা কী করে মাপবে? তার জন্যে $\vec{w}$ বরাবর একটা লাইন কল্পনা করো, এটা যেন একটা পর্দা (Fig 9)। এই পর্দার উপরে দুদিক থেকেই আলো পড়ছে লম্বভাবে, ফলে বিন্দুগুলোর ছায়া পড়েছে পর্দার উপরে। বোঝার সুবিধার জন্য ছায়াগুলোকে কালো রঙে, এবং মূল বিন্দুগুলোকে হাল্কা করে দেখিয়েছি। Fig 10-এও একই কাজ করেছি, কিন্তু দিকটা অন্যভাবে নিয়েছি, ফলে পর্দার অবস্থান ঘুরে গেছে। লক্ষ্য করো, Fig 9-এর বেলায় ছায়াগুলো যতটা ছড়ানো ছেটানো ছিল, Fig 10-এর বেলায় তার চেয়ে অনেকটাই বেশী ঘন হয়ে আছে। এই ছায়াগুলো হল একটা লাইন বরাবর বিন্দু, তাই ওরা কতটা ছড়ানো ছেটানো, সেটা variance দিয়ে মাপা যাবে। এইভাবে যেকোনো $\vec{w}$-এর জন্যই আমরা একটা করে variance পাব। এবার ব্যাপারটা একটু অংকের ভাষায় গুছিয়ে লিখে নিই। প্রথমে point-গুলোকে একেকটা vector বলে ভাবো (Fig 11), তবে তার ছায়ার অবস্থানটা আসলে হবে $\vec{w}$ বরাবর তার component. যেমন,

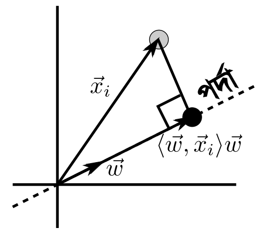


Fig 11

বিন্দুটা $\vec{x}_i$ হলে component-টা হবে

$$\langle\vec{x}_i,\vec{w}\rangle\,\vec{w}.$$

Fig 11 দেখো। এই component-গুলো সব একই লাইন বরাবর আছে ($\vec{w}$-এর লাইন বরাবর)। তাই যদি $z_i = \langle\vec{x}_i,\vec{w}\rangle$ নিই, তবে মনে করতে পারো যেন এই লাইনটাই আমাদের $\mathbb{R}$ আর component-টা হল সেই লাইন বরাবর একটা সংখ্যা, z_i. এদেরই variance আমরা বার করব--

$$\frac{1}{n}\sum_1^n (z_i-\overline{z})^2. \qquad (*)$$

এই সংখ্যাটা মাপছে $\vec{w}$ দিক বরাবর point-গুলো কতটা ছড়িয়ে-ছিটিয়ে আছে। এই পরিমাণটা দুটো জিনিসের উপর নির্ভর করছে--এক, point-গুলো কীভাবে আছে, তার উপর, এবং দুই, $\vec{w}$ কীভাবে নেওয়া হয়েছে, তার উপর। আমরা চাই (*)-কে এমনভাবে সাজিয়ে লিখতে, যাতে এই দুটো প্রভাবকে আলাদা করে বোঝা যায়। এর জন্য লক্ষ্য কর যে

$$\overline{z} = \frac{1}{n}\sum_1^n z_i = \frac{1}{n}\sum_1^n \langle\vec{x}_i,\vec{w}\rangle = \left\langle\overline{\vec{x}},\vec{w}\right\rangle.$$

সুতরাং

$$z_i - \overline{z} = \left\langle\vec{x}_i - \overline{\vec{x}},\vec{w}\right\rangle.$$

আমরা এখানে Euclidean inner product নিয়ে কাজ করছি। তাই দুটো vector $\vec{u},\vec{v}$-এর মধ্যে inner product-টাকে এই দুইভাবেই লিখতে পারি--

$$\langle\vec{u},\vec{v}\rangle = \vec{v}'\vec{u} = \vec{u}'\vec{v}.$$

সুতরাং--

$$\begin{aligned}
\frac{1}{n}\sum_1^n (z_i-\overline{z})^2 &= \frac{1}{n}\sum_1^n \left\langle\vec{x}_i-\overline{\vec{x}},\vec{w}\right\rangle\cdot\left\langle\vec{x}_i-\overline{\vec{x}},\vec{w}\right\rangle \\
&\quad \text{এবার আমরা একই inner product-কে দুভাবে লিখব--} \\
&= \frac{1}{n}\sum_1^n \left[\vec{w}'(\vec{x}_i-\overline{\vec{x}})\cdot(\vec{x}_i-\overline{\vec{x}})'\vec{w}\right] \\
&= \vec{w}'\left[\frac{1}{n}\sum_1^n(\vec{x}_i-\overline{\vec{x}})(\vec{x}_i-\overline{\vec{x}})'\right]\vec{w} \\
&= \vec{w}'V\vec{w},
\end{aligned}$$

যেখানে

$$V = \frac{1}{n}\sum_1^n(\vec{x}_i-\overline{\vec{x}})(\vec{x}_i-\overline{\vec{x}})'.$$

লক্ষ্য কর যে, এটা একটা quadratic form. এভাবে লেখার সুবিধাটা হল এখানে V কেবলমাত্র point-গুলোর উপরই নির্ভর করে, $\vec{w}$-এর উপরে নয়। তাই আমরা বলতে পারি যে, বিভিন্ন দিক বরাবর point-গুলো কতটা করে ছড়িয়ে-ছিটিয়ে আছে, তার সম্পূর্ণ হদিশ রয়েছে এই matrix-টার মধ্যে। তোমার যে $\vec{w}$-টা প্রয়োজন সেইটা দিয়ে quadratic form বানিয়ে নিলেই সেই দিকের উত্তরটা বেরিয়ে আসবে। এই V-টা statistics-এ একটা অত্যন্ত গুরুত্বপূর্ণ জিনিস, একে বলে **variance-covariance matrix.**

Exercise 6: ধরো তোমাকে কিছু লোকের উচ্চতা এবং ওজন দিলাম--

উচ্চতা	(cm)	175	176	173
ওজন	(kg)	70	76	64

এর ভিত্তিতে variance-covariance matrix বার করো।
HINT: এরকম অংক statistician-দের হরদম করতে হয়। যেকোনো statistical software দিয়ে কাজটা মুহূর্তের মধ্যে হয়ে যাবে। কিন্তু অংকটা কম্পিউটার ছাড়া করতে খুব মজা নাও লাগতে পারে। ক্যালকুলেটর হাতের কাছে রেখো। কী করে শুরু করবে বলি। এখানে তিনটে মানুষের সম্পর্কে তথ্য আছে, তাই $n = 3$. প্রত্যেকটা মানুষের তথ্যকে আমরা একটা vector হিসেবে লিখব, $\vec{x} =$ (উচ্চতা, ওজন). সেই হিসেবে $\vec{x}_1 = (175, 70)$, ইত্যাদি। এবার variance-covariance matrix বার করার ফর্মুলাটা লাগালেই হবে। ■

55.2 Physics

Quadratic form ব্যাপারটা অনেক অপ্রত্যাশিত জায়গায় কাজে লাগে। একটা উদাহরণ দিই physics থেকে। পরে যদি কখনো rigid motion নিয়ে পড়াশোনা কর এই উদাহরণটা কাজে লাগবে। এই উদাহরণটা বোঝার জন্য সামান্য কিছু প্রস্তুতি চাই--

- যখন কোনো জিনিস শূন্যে circular (বৃত্তাকার) পথে ঘোরে (Fig 12), তখন সেই circle-এর centre দিয়ে আমরা circle-এর তলের সঙ্গে লম্বভাবে একটা লাইন টানতে পারি, একে বলে **rotation axis**.
- জিনিসটার ঘূর্ণনটাকে দুভাবে বোঝানো যায়, প্রাথমিক অবস্থান থেকে t সময়ে কত কোণ $\theta(t)$ ঘুরেছে সেটা দিয়ে, বা কতটা দূরত্ব circle-এর পরিধি বরাবর গেছে, সেই $s(t)$ দিয়ে।
- সেই অনুযায়ী differentiate করে ওর speed-ও দুভাবে বার করা যায়, $w(t) = \theta'(t)$ হল **angular speed,** আর $v(t) = s'(t)$ হল **linear speed.**
- যদি circle-টার radius হয় r, তবে এদের মধ্যে সহজ সম্পর্ক[2] আছে, $s(t) = r\,\theta(t)$ এবং তাই $v(t) = r\,w(t)$.
- আমরা ঘূর্ণনটাকে Fig 13-এর মত একটা vector দিয়ে প্রকাশ করতে পারি, $\vec{w}$, যেটা rotation axis বরাবর রয়েছে, অভিমুখটা ঠিক হচ্ছে right hand rule[3] দিয়ে, এবং যার দৈর্ঘ্য হল angular speed-এর সমান।

[2] এর জন্য কোণটা মাপতে হবে radian এককে।
[3] বুড়ো আঙুল ছাড়া ডানহাতের বাকি আঙুলগুলো ঘূর্ণনের দিকে বাঁকালে বুড়ো আঙুল যেদিক নির্দেশ করে, সেই অভিমুখটা।

Fig 12

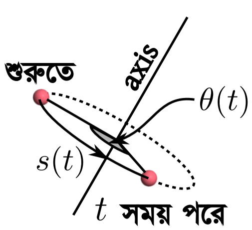


Fig 13

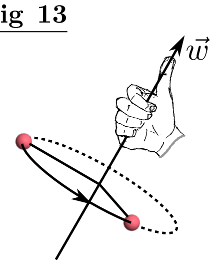


Fig 14

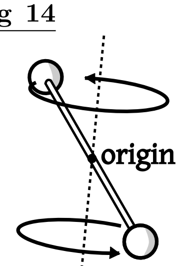
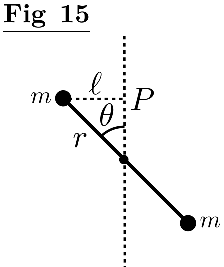


Fig 15



- জিনিসটার mass যদি m হয় আর linear speed যদি কোনো মুহূর্তে v হয়, তবে সেই মুহূর্তে তার মধ্যে kinetic energy-র পরিমাণ হবে $\frac{1}{2}mv^2$.

এবার ধরো একটা জিনিস আছে Fig 14-এর মত, দুটো ছোটো কিন্তু ভারী লোহার বল একটা সরু ডাণ্ডার দুপ্রান্তে ডাম্বেলের মত করে লাগানো। বা ভাবতে পারো যেন একটা কাঠির দু মাথায় দুটো আলুর দম। দুটো আলুই সমান ভারী, কাঠিটা খুবই সরু, তার mass নগণ্য। আমরা origin-টা নিয়েছি কাঠির মধ্যবিন্দুতে। সেই বিন্দু দিয়ে যা খুশি একটা axis বরাবর ডাম্বেলটা ঘুরছে। এই ঘূর্ণনটাকে আমরা যথারীতি একটা vector দিয়ে লিখব--

$$\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}.$$

এই ঘোরার ফলে লোহার বল দুটোও ঘুরছে। পুরো জিনিসটার মধ্যে মোট কত kinetic energy আছে, আমরা তা বার করব। এই kinetic energy-টা rotation axis-টার অবস্থানের উপর নির্ভর করে। সেটা প্রথমে সহজ বুদ্ধি দিয়ে অনুভব করা যাক--

- ধরো axis-টা নিয়েছি ডাম্বেলটার সঙ্গে সমকোণে। এবার ঘোরাটা হবে একেবারে ফ্যানের মত। এখানে kinetic energy প্রচুর।

- কিন্তু যদি axis-টা নিতাম ডাণ্ডাটা বরাবর, তবে ঘোরাটা হত স্ক্রুয়ের মত। বল দুটো সাইজে ছোটো বলেছি, তাই সেই ঘোরা চোখে ধরাই পড়বে না, kinetic energy-ও হবে নগণ্য।

তবেই বুঝলে rotation axis বদলালে kinetic energy-ও বদলে যাচ্ছে। এবার ব্যাপারটা অংক কষে বার করা যাক। এর জন্য কিছু মাপজোখ দরকার, যেগুলো দেখিয়েছি Fig 15-এ। প্রত্যেকটা বলই circular পথে ঘুরবে, যার radius হল $r\sin\theta$. এদিকে angular speed হল $\|\vec{w}\|$. তাই linear speed হবে $r\sin\theta\,\|\vec{w}\|$. সুতরাং প্রত্যেকটা বলের kinetic energy হচ্ছে $\frac{1}{2}mr^2\sin^2\theta\,\|\vec{w}\|^2$. যেহেতু দুটো বল আছে, তাই মোট kinetic energy হচ্ছে--

$$mr^2\sin^2\theta\,\|\vec{w}\|^2.$$

এটা মোটেই জটিল কিছু নয়, কিন্তু একটাই সমস্যা--এর মধ্যে ডাম্বেলটার ভূমিকা আর rotation vector-এর ভূমিকা একাকার হয়ে আছে, ওদেরকে আলাদা করে বোঝা যাচ্ছে না, কারণ ওই θ বার করতে ডাম্বেলের ডাণ্ডা আর rotation axis দুটোই একইসঙ্গে দরকার। আমরা kinetic energy-টাকে এমনভাবে লিখতে চাই, যাতে এদের ভূমিকা আলাদা করে বোঝা যায়। তার জন্য প্রথমে খালি একটা বল নিয়ে কাজ করি। ধরো কোনো মুহূর্ত t-তে এই বলটার অবস্থান হল--

$$\vec{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

এরা সকলেই আসলে t-এর function, যদিও দেখতে জটিল হয়ে যাবে বলে $x(t)$, $y(t)$ ইত্যাদি লিখিনি। লক্ষ্য করো, এখানে $\|\vec{r}\| = r$.
তাহলে axis থেকে এর দূরত্বটা হবে--

$$\ell^2 = \|\vec{r}\|^2 - \left\|\frac{\langle\vec{r},\vec{w}\rangle}{\|\vec{w}\|^2}\vec{w}\right\|^2 = \|\vec{r}\|^2 - \frac{\langle\vec{r},\vec{w}\rangle^2}{\|\vec{w}\|^2}.$$

এইটা দেখে ভয় পেয়ে যেও না। এটা আমাদের পূর্বপরিচিত। আমরা কদিন আগে একটা vector বরাবর আরেকটা vector-এর component বার করা শিখেছিলাম। এটা সেই জিনিসটারই একটা প্রয়োগ--

> $\vec{w}$ বরাবর $\vec{r}$-এর component হল $\frac{\langle \vec{r},\vec{w}\rangle}{\langle \vec{w},\vec{w}\rangle}\vec{w}$. তার মানে $\vec{r}$-এর ডগা থেকে $\vec{w}$-এর লাইনের উপর একটা লম্ব টানলে লম্বটা $\vec{w}$-এর লাইনকে এইখানে ছেদ করে (Fig 15-এ P বিন্দু)। সুতরাং $\vec{w}$-এর লাইনের থেকে $\vec{r}$-এর দূরত্ব হল $\vec{r}$-এর ডগা থেকে এই বিন্দুটার দূরত্বটার দূরত্ব, যেটাকে ℓ নাম দিয়েছি। এবার খালি একবার Pythagoras লাগালেই ℓ^2-এর এই ফর্মুলাটা পেয়ে যাবে।

এখানে $\|\vec{w}\|$ হল angular speed. তাই বলটার linear speed হবে $v = \ell\|\vec{w}\|$. সুতরাং kinetic energy হবে $\frac{1}{2}mv^2$. এবার এটাকে আমরা matrix-এর কায়দায় কী করে সাজিয়ে লিখব দেখো।

$$\begin{aligned}
\frac{1}{2}mv^2 &= \frac{1}{2}m\ell^2\,\|\vec{w}\|^2\\
&= \frac{1}{2}m\left[\|\vec{r}\|^2 - \frac{\langle \vec{r},\vec{w}\rangle^2}{\|\vec{w}\|^2}\right]\|\vec{w}\|^2\\
&= \frac{1}{2}m\left[\|\vec{r}\|^2\,\|\vec{w}\|^2 - \langle \vec{r},\vec{w}\rangle^2\right]\\
&= \frac{1}{2}m\left[\|\vec{r}\|^2\,\vec{w}'\vec{w} - (\vec{r}'\vec{w})(\vec{r}'\ \vec{w})\right]\\
&= \frac{1}{2}m\left[\|\vec{r}\|^2\,\vec{w}'\vec{w} - (\vec{w}'\vec{r})(\vec{r}'\ \vec{w})\right]\\
&= \vec{w}'\underbrace{\left[\frac{1}{2}m\left\{\|\vec{r}\|^2 I - \vec{r}\vec{r}'\right\}\right]}_{M}\vec{w}.
\end{aligned}$$

এরকম দুটো বল আছে, তাই মোট kinetic energy হবে--

$$m\vec{w}'\underbrace{\left[\left\{\|\vec{r}\|^2 I - \vec{r}\vec{r}'\right\}\right]}_{M}\vec{w}.$$

লক্ষ্য কর যে, এটা আসলে $\vec{w}$-এর একটা quadratic form! এই যে মাঝখানের matrix-টা, যাকে M নাম দিয়েছি, এর একটা জাঁদরেল নাম[4] আছে--**moment of inertia matrix.** কখনো বা একে বলে **inertia tensor.** লক্ষ্য করো এই M-টা পুরোটাই নির্ভর করছে কেবলমাত্র ডাম্বেলটার উপরে, $\vec{w}$-এর উপরে নয়। সুতরাং quadratic form হিসেবে লেখার ফলে আমরা kinetic energy-র মধ্যে ডাম্বেলের প্রভাবটাকে আলাদা করে প্রকাশ করতে পারলাম। অবশ্য মনে রেখো যে, M কিন্তু t-এর function. সেটা একটু অদ্ভুত লাগতে পারে, কারণ ডাম্বেলটা তো সময়ের সাথে সাথে বদলাচ্ছে না! আসলে যখন বলছি যে, M-টা ডাম্বেলটার উপর নির্ভর করে, আমরা বোঝাতে চাইছি যে, সেটা ডাম্বেলটার গঠন এবং অবস্থান দুয়ের উপরেই নির্ভর করে। যেহেতু ডাম্বেলটার অবস্থান সময়ের সাথে সাথে বদলাচ্ছে, তাই M-ও t-এর function হচ্ছে। এই matrix-টা physics-এ খুব গুরুত্বপূর্ণ স্থান অধিকার করে আছে।

Exercise 7: Fig 16-এ একটা বল রয়েছে y-axis-এর উপরে। এর moment of inertia matrix বার করো। ■

55.3 Multivariable calculus

Real analysis-এ আমরা Taylor polynomial-এর কথা শিখি। ধরো কোনো function $f : \mathbb{R} \to \mathbb{R}$ দেওয়া আছে, আর দেওয়া আছে কোনো $a \in \mathbb{R}$. আমাদের কাজ হল এমন একটা polynomial বার করা, যেটা ওই a-র কাছাকাছি যতটা সম্ভব $f(x)$-এর মত হয়। এরকম polynomial বার করতে চাওয়ার পিছনের মূল কারণ হল polynomial-রা বেশ ভদ্রসভ্য function, ওদের নিয়ে কাজ করা সহজ। কিন্তু f-টা বেশ বদখত কিছু হতেই পারে। একটা বদখত জিনিসকে যদি

[4] এই নামগুলো তখনই ব্যবহার করা যায়, যখন origin-টাকে নেওয়া হয় ডাম্বেলটার centre of mass-এ (মানে ভরকেন্দ্রে)। আমরা এখানে সেটাই নিয়েছিলাম।

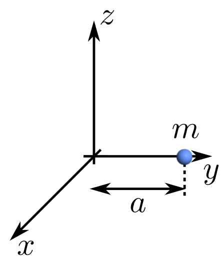


Fig 16

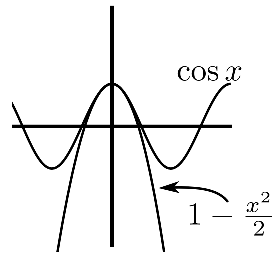


Fig 17

একটা ভদ্রসভ্য জিনিস দিয়ে approximate করা যায়, তবে কাজের সুবিধা। এখানে approximation বলতে কী বোঝাচ্ছি? বোঝাচ্ছি এমন একটা polynomial, ধরো $p(x)$, যার degree যদি k হয়, তবে যেন $x = a$-তে $f(x)$ আর $p(x)$-এর value তো সমান চাইই, সেই সঙ্গে যেন আবার প্রথম k-খানা derivative-ও মিলে যায়। অবশ্য তার আগে এটা ধরে নিচ্ছি যে, $f(x)$-টা ওই বিন্দুতে অন্ততঃ k-বার differentiable. ঠিক কীভাবে polynomial-টা নিলে সেটা হবে? এরই উত্তর হল $x = a$-তে $f(x)$-এর Taylor polynomial, যার সংজ্ঞা--

$$p(x) = f(a) + \sum_{i=1}^{k} \frac{f^{(i)}(a)}{i!}(x-a)^i.$$

এর মধ্যে সবচেয়ে গুরুত্বপূর্ণ কেসটা হল যখন $k = 1$. সেক্ষেত্রে $p(x)$-এর গ্রাফটা একটা সরলরেখা, যাকে বলে ওই বিন্দুতে $f(x)$-এর tangent. এটাকে ভাবা যেতে পারে সেই সরলরেখা, যেটা ওই বিন্দুর কাছাকাছি $f(x)$-এর সবচেয়ে ভালো linear approximation. একইভাবে যদি $k = 2$ নিই, তবে পাব সবচেয়ে ভালো quadratic approximation. ফর্মুলাটা হল--

$$p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2.$$

এর ডানদিকে যে তিনটে term রয়েছে, তাদেরকে আমরা constant, linear এবং quadratic আখ্যা দিতে পারি।

Example 4: Fig 17-এ আমরা $\cos x$-এর quadratic approximation দেখিয়েছি $x = 0$-তে--

$$p(x) = 1 - \frac{x^2}{2}.$$

অবশ্যই quadratic approximation-টা একটা parabola, এবং $\cos x$-এর গ্রাফটা মোটেই একটা parabola নয়। কিন্তু তাও লক্ষ্য করো, $x = 0$-র কাছে দুটো গ্রাফ যেন একেবারে পরস্পরের সঙ্গে মিশে গেছে। ■

এ যাবৎ আমরা $f : \mathbb{R} \to \mathbb{R}$ নিয়ে কাজ করছিলাম। কিন্তু একইরকম কাজ করা যেতে পারে $f : \mathbb{R}^n \to \mathbb{R}$-এর জন্যও। সেক্ষেত্রেও তিনটে term থাকবে, এবং ওই quadratic term-টার ভূমিকা পালন করার জন্য quadratic form-এর ডাক পড়বে। এক্ষেত্রে polynomial-টার চেহারা হবে এইরকম--

$$p(x) = c + \sum_i \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i) + \frac{1}{2}\sum_i\sum_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})(x_i - a_i)(x_j - a_j). \qquad (*)$$

Example 5: ধরো $f(x, y) = \sin x \cos y$ নিলাম। তবে $(x, y) = \left(\frac{\pi}{4}, \frac{\pi}{4}\right)$ বিন্দুতে (*) কী হবে?

SOLUTION: কষলেই দেখবে উত্তর আসছে

$$\frac{1}{2}\left(1 + x - y - \left(x + y - \tfrac{\pi}{2}\right)^2\right).$$

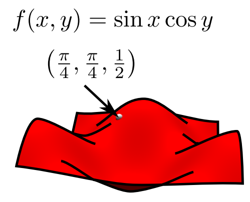


Fig 18

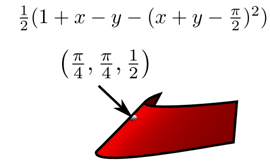


Fig 19

আমরা $f(x,y)$-এর গ্রাফ এঁকে দেখিয়েছি Fig 18-এ, আর এই quadratic approximation-টার ছবি দিয়েছি Fig 19-এ। দেখতে দুটো খুবই আলাদা, কিন্তু ওই $(x,y) = \left(\frac{\pi}{4}, \frac{\pi}{4}\right)$ বিন্দুর ধারেকাছে ওরা সত্যিই একইরকম দেখতে। ■

একটু চিন্তা করলেই বুঝবে যে, (*)-এর ডানদিকের তৃতীয় term-টা আসলে একটা quadratic form, যাকে matrix দিয়ে লিখলে হবে $\frac{1}{2}(\vec{x}-\vec{a})'H(\vec{x}-\vec{a})$, যেখানে H-এর (i,j)-th entry হল

$$h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}).$$

এই symmetric matrix-টাকে বলে $\vec{x} = \vec{a}$-তে $f(\vec{x})$-এর **Hessian**. একটা উদাহরণ করে গা গরম করে নেওয়া যাক।

Example 6: Find the Hessian matrix of the function $f(x,y) = x^3 - 2xy - y^6$ at $(1,2)$.[2] **(2022.CC12.3di)**

SOLUTION:

We have $\frac{\partial f}{\partial x} = 3x^2 - 2y$ and $\frac{\partial f}{\partial y} = -2x - 6y^5$.

Hence

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= 6x, \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial^2 f}{\partial y \partial x} = -2, \\ \frac{\partial^2 f}{\partial y^2} &= -30y^4. \end{aligned}$$

Evaluating these at $(x,y) = (1,2)$, the required Hessian matrix is

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 6 & -2 \\ -2 & -480 \end{bmatrix}.$$

■

Exercise 8: Find the Hessian matrix of the function $f(x,y,z) = e^x - xy + x^2 z$ at $(1,2,3)$. ■

Quadratic approximation-এর অংশ হিসেবে কী করে Hessian matrix আসে, সেটা দেখলাম। Hessian দিয়ে আর কোনো উপকার হয়? এর সংক্ষিপ্ত উত্তর হল $f : \mathbb{R} \to \mathbb{R}$-এর বেলায় second derivative দিয়ে যা কাজ হয়, $f : \mathbb{R}^n \to \mathbb{R}$-এর বেলায় সেই একই কাজ হয় Hessian দিয়ে। আরো খুঁটিনাটি বোঝার জন্য quadratic form-দের নিয়ে আরেকটু জানা থাকা দরকার। সেটা শেখার পর আমরা আবার Hessian-এর আলোচনায় ফিরব।

DAY 56
Congruence (part 1)

56.1 ব্যাপারটা কী?

একটা quadratic form দিলেই তুমি তা থেকে একটা symmetric matrix বার করতে পারো। এটা আমরা শিখেছি। কিন্তু লক্ষ্য করেছ কি যে, এর মধ্যে একটা basis-এর গল্প লুকিয়ে আছে? নীচের উদাহরণটা দেখলেই সেটা বুঝবে।

Example 7: ধরো $f : \mathbb{R}^2 \to \mathbb{R}$ একটা quadratic form, যেখানে $f(u,v) = u^2 + 2v^2 + 6uv$. এবার তোমাকে $\mathbb{R}^2$-র মধ্যে Fig 20-এ দেখানো vector-টা দিলাম। এর জন্য f-এর value কী হবে?

SOLUTION: এটা বার করা অসম্ভব। কারণ f বার করার জন্য vector-টাকে আগে (u,v) আকারে লিখে নিতে হবে। এবং তার মানেই হল কোনো basis ঠিক করে নিয়ে তার সাপেক্ষে vector-টাকে প্রকাশ করা। যেহেতু ছবিতে কোনো basis দেওয়া নেই, তাই সে কাজটা করা যাচ্ছে না। যদি Fig 21-এর মত basis নিই $\{(1,0), (0,1)\}$, তবে (u,v) হচ্ছে $(1,1)$, কিন্তু যদি Fig 22-এর basis-টা নিতাম $\{(0,1), (-1,0)\}$, তবে বেরোত $(u,v) = (1,-1)$. বুঝতেই পারছ অন্ধের মত f-এর একই ফর্মুলা দুইক্ষেত্রেই ব্যবহার করলে f-এর value একই vector-এর জন্য দুরকম আসবে। সেটা তো আর হতে পারে না, কারণ হাজার হোক, $f : \mathbb{R}^2 \to \mathbb{R}$ তো একটা function! ■

এই উদাহরণ থেকেই বুঝতে পারছ যে, যখন আমরা কোনো quadratic form-কে ফর্মুলা দিয়ে লিখছি, তখন সেটা কোনো একটা basis-এর সাপেক্ষে লিখছি। Basis-টা বদলালে ফর্মুলাটাকেও বদলাতে হবে। যেহেতু ফর্মুলাটাকেই সাজিয়ে লিখে matrix-টা তৈরী হয়, তাই কোনো quadratic form-এর matrix-টাও basis-এর উপর নির্ভর করে। Basis বদলালে matrix-ও বদলাবে। এতক্ষণ যে আমাদের আলোচনায় basis-এর প্রসঙ্গ আসে নি, তার কারণ আমরা খালি ফর্মুলাটাকে সাজিয়ে matrix বানানো শিখছিলাম। পুরো কাজটা একটা basis নিয়েই হচ্ছিল বলে basis-টাকে স্পষ্ট করে উল্লেখ করতে হয়নি।

এই basis-নির্ভরতাটা আমরা ইতিমধ্যেই linear transformation-এর বেলাতেও দেখেছিলাম। সেখানেও একই linear transformation-কে বিভিন্ন basis-এর সাপেক্ষে প্রকাশ করতে গেলে আলাদা আলাদা matrix আসত।

Example 8: Canonical basis-এর সাপেক্ষে একটা quadratic form-এর matrix হল $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. তবে $\{(1,1), (1,-1)\}$-এর সাপেক্ষে এর matrix কী হবে?

Fig 20



Fig 21

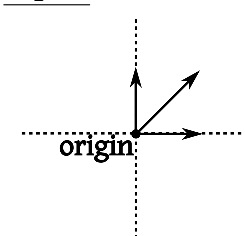
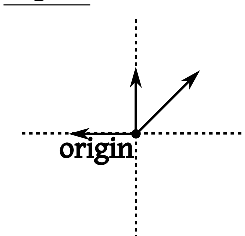


Fig 22



SOLUTION: ধরো একটা vector আছে, যেটা $\{(1,1),(1,-1)\}$-এর সাপেক্ষে (u,v). তার মানে canonical basis-এর সাপেক্ষে vector-টা হল

$$u(1,1)+v(1,-1)=(u+v,u-v).$$

সুতরাং quadratic form-টা হল

$$\begin{aligned}\begin{bmatrix} u+v & u-v \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} u+v \\ u-v \end{bmatrix} &= (u+v)^2+3(u-v)^2+4(u+v)(u-v) \\ &= \cdots = 8u^2-4uv = \begin{bmatrix} u & v \end{bmatrix}\begin{bmatrix} 8 & -2 \\ -2 & 0 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}.\end{aligned}$$

সুতরাং উত্তর হল $\begin{bmatrix} 8 & -2 \\ -2 & 0 \end{bmatrix}$. ■

এই অংকে আমরা matrix-টা গায়ের জোরে বার করলাম। যদি dimension একটু বেশী হত, তবে এভাবে এগোনো শক্ত। কিন্তু change of basis matrix ব্যবহার করলে সবটা বেশ গুছিয়ে লেখা যায়। সেই জিনিসটা কী ছিল একটু মনে করিয়ে দিই। ধরো দুটো basis আছে $\{\vec{u}_1,...,\vec{u}_n\}$ এবং $\{\vec{v}_1,...,\vec{v}_n\}$. এবার একটা vector দিলাম $\vec{x}$ যাকে প্রথম basis-টার সাপেক্ষে লিখলে হয় $(a_1,...,a_n)$ এবং দ্বিতীয়টার সাপেক্ষে লিখলে হয় $(b_1,...,b_n)$, মানে

$$\vec{x}=\sum_1^n a_i\vec{u}_i=\sum_1^n b_i\vec{v}_i.$$

আমাদের উদ্দেশ্য হল এদের মধ্যে একটা দেওয়া থাকলে অন্যটা বার করা, ধরো a_i-গুলো দেওয়া আছে, আর আমরা তা থেকে b_i-গুলো বার করতে চাই। সেই কাজটা একটা matrix গুণ করে করা যায়, মানে এমন একটা matrix পাওয়া যায়, C, যাতে--

$$\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}=C\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

হবে। এই C-টা খালি basis দুটোর উপরই নির্ভর করবে, $\vec{x}$-এর উপরে নয়, অর্থাৎ যেকোনো $\vec{x}$-এর জন্যই একই matrix দিয়ে কাজ চলবে। এই C-কেই বলে একটা change of basis matrix. এটা বার করার কায়দা হল প্রথমে u_i-গুলোকে $\{\vec{v}_1,...,\vec{v}_n\}$-এর সাপেক্ষে tuple হিসেবে লিখে নেওয়া, এবং সেই tuple-গুলোকে পাশাপাশি column-এর মত সাজিয়ে C বানানো। Change of basis matrix-রা সর্বদাই nonsingular matrix হয়, এবং বিপরীতপক্ষে যেকোনো nonsingular matrix-ই একটা change of basis matrix-এর ভূমিকা পালন করতে পারে (উপযুক্ত একজোড়া basis-এর জন্য)। এবার ফের আমাদের মূল আলোচনায় ফেরা যাক।

Example 9: আগের অংকটাই আবার করো, কিন্তু change of basis matrix ব্যবহার করে।

SOLUTION: আগের বার আমরা দেখেছিলাম যে নতুন basis-এর সাপেক্ষে যে vector-টা (u,v), সেটাই canonical basis-র সাপেক্ষে $(u+v,u-v)$. আমরা লিখতে পারি

$$\begin{bmatrix} u+v \\ u-v \end{bmatrix}=\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{P}\begin{bmatrix} u \\ v \end{bmatrix}.$$

ওই matrix-টাই হল এখানে আমাদের change of basis matrix. সুতরাং আমাদের quadratic form-টা হচ্ছে

$$\left(P\begin{bmatrix} u \\ v \end{bmatrix}\right)'\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\left(P\begin{bmatrix} u \\ v \end{bmatrix}\right)=\begin{bmatrix} u & v \end{bmatrix}P'\underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}}_{A}P\begin{bmatrix} u \\ v \end{bmatrix}.$$

তার মানে মূল matrix-টাকে যদি A বলি, তবে নতুন matrix-টা হল $P'AP$. ■

Exercise 9: Canonical basis-এর সাপেক্ষে একটা quadratic form-এর matrix হল $A = \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}$. তবে $\{(1,1), (1,-1)\}$-এর সাপেক্ষে এর matrix কী হবে? ■

এখানে থেকেই আসে নীচের সংজ্ঞাটা--

> DEFINITION: **Congruent**
>
> Two symmetric matrices, A and B, are called **congruent** if $B = P'AP$ for some nonsingular matrix, P.

এটা দেখা শক্ত নয় যে, congruence হল একটা equivalence. দুটো symmetric matrix যদি পরস্পরের সঙ্গে congruent হয়, তার মানে ওরা আসলে একই quadratic form-কে প্রকাশ করছে, খালি হয়তো অন্য basis-এর সাপেক্ষে। স্বভাবতঃই দুটো symmetric matrix হাতে পেলেই আমরা জানতে চাই ওরা পরস্পরের সঙ্গে congruent কিনা। ব্যাপারটা অনেকটা ছোটোবেলায় শেখা ভগ্নাংশের মত। যদি তোমাকে $\frac{3}{6}$ এবং $\frac{50}{100}$ দিই, তবে ওদের দেখতে আলাদা হলেও ওরা আসলে ভগ্নাংশ হিসেবে একই। এই ব্যাপারটা বোঝার সহজতম কায়দা ছিল "কাটাকাটি" করে দেখা। দুটোকেই কাটাকাটি করলে একই জিনিস পাবে, $\frac{1}{2}$. অমনি বলতে পারব যে, $\frac{3}{6}$ আর $\frac{50}{100}$ আসলে একই ভগ্নাংশকে প্রকাশ করছে। এই কাটাকাটির কায়দাটা কাজ করে কেন? দুটো কারণে--

- প্রথমতঃ, একটা ভগ্নাংশকে $\frac{m}{n}$ রূপে অসংখ্যভাবে লেখা যায় বটে, কিন্তু যদি $gcd(m,n) = 1$ এবং $n > 0$ শর্ত চাপাই, তবে ঠিক একভাবেই লেখা যায়। এই বিশেষ রূপটাকে বলে ভগ্নাংশটার **reduced** রূপ।
- দ্বিতীয়তঃ, একটা ভগ্নাংশকে যে রূপেই দেওয়া হোক না কেন, "কাটাকাটি" করে সেটাকে reduced রূপে এনে ফেলা যায়।

এই ধারণাটা শুধু ভগ্নাংশের ক্ষেত্রেই নয়, অংকের দুনিয়ায় আরো অনেক জায়গাতেই কাজে লাগে। আসলে অংকের জগতে আমাদের কারবার abstract সব ব্যাপার-স্যাপার নিয়ে। কিন্তু সমস্যা হল abstract জিনিসপত্র লিখে বোঝানো কঠিন। তাই অংকে হরেকরকমের কায়দা বার করা হয়েছে সেগুলোকে লিখে প্রকাশ করতে। কিন্তু প্রায়শঃই দেখা যায় যে, সেই সব কায়দায় একই abstract ধারণা একাধিকভাবে লেখা সম্ভব। সুতরাং খালি লেখাটুকু আলাদা দেখেই ওদেরকে আলাদা জিনিস বলে ভেবে ফেললে মুস্কিল। এই সমস্যা থেকেই equivalence-এর উৎপত্তি। যেসব জিনিস একই abstract ধারণাকে প্রকাশ করে, সেগুলো মিলে একটা equivalence class হয়। এবার এমন একটা বিশেষ চেহারা গণিতজ্ঞরা খুঁজে পেতে চান, যাতে যেকোনো equivalence class-এর মধ্যেই সেই চেহারার ঠিক একজনই প্রতিনিধি থাকবে, এবং এমন কোনো একটা কায়দা থাকবে যার সাহায্যে সেই equivalence class-এর যেকোনো সদস্যকেই ধাপে ধাপে সেই প্রতিনিধিতে পরিণত করা যাবে। যদি এরকম কোনো চেহারা পাওয়া যায়, তবে তাকে সাধারণতঃ সেই equivalence-এর জন্য **reduced form** বা **normal form** বা **canonical form** বলা হয়। আমরা এবার congruence-এর জন্য এরকম একটা রূপ শিখব।

56.2 Congruence-এর জন্য normal/canonical form

Congruence-এর বেলায় canonical form বার করার জন্য যে কায়দা, তার পিছনের মূল ধারণাটা আমরা আগে থেকেই জানি--square complete করা। চট্ করে মনে করিয়ে দিই।

56.2.1 Square complete করা

হায়ার সেকণ্ডারীর সময়ে আমরা শিখি--

- $\int \frac{dx}{\sqrt{a^2-x^2}} = \sin^{-1}\frac{x}{a} + c$ (যেখানে $a > 0$),
- $\int \frac{dx}{\sqrt{a^2+x^2}} = \sinh^{-1}\frac{x}{a} + c$ (যেখানে $a > 0$),
- $\int \frac{dx}{\sqrt{x^2-a^2}} = \log\left(2(\sqrt{x^2-a^2}+x)\right) + c.$

এবার ধরো তোমাকে এই integral-টা বার করতে দিলাম--

$$\int \frac{dx}{-9x^2+12x+25}.$$

তবে তুমি এইভাবে এগোবে--

$$\begin{aligned} -9x^2+12x+25 &= -(3x)^2 + 2\times(3x)\times 2 - 2^2 + 2^2 + 25 \\ &= -(3x-2)^2 + \left(\sqrt{29}\right)^2. \end{aligned}$$

এরপর তুমি $y = 3x - 2$ আর $a = \sqrt{29}$ নেবে তাহলে integral-টা প্রথম চেহারাটার মত হয়ে যাবে। যেভাবে $-9x^2+12x+25$-কে দুটো square-এ ভেঙে লিখলাম, এটাই হল square complete করা। এর জন্য প্রধান হাতিয়ার অবশ্যই এই ফর্মুলাটা--

$$(a+b)^2 = a^2 + b^2 + 2ab$$

আমরা প্রথমে $-9x^2$ দেখে মাইনাসটা আলাদা করে নিলাম, তারপর $9x^2$-কে $(3x)^2$ হিসেবে লিখলাম। সুতরাং আমাদের $(a+b)^2$-এর জন্য $a = 3x$ নেব। এবার আছে $12x$ যেটাকে $2ab$ বানাতে হবে, সুতরাং $b = 2$ নিতে হচ্ছে। $(a+b)^2$-এ একটা b^2 চাই, তাই b^2-টা যোগ বিয়োগ করে দিলাম।

এখানে একটাই variable ছিল, x. কিন্তু এই কায়দাটা সব সময়েই করা যায়, $x, y, z, \ldots$ যতগুলো খুশী variable থাক না কেন। যেমন ধরো এক্ষুণি যে অংকটা করলাম সেটাকেই x, y দিয়ে লেখা যায়--

$$-9x^2 + 12xy + 25y^2 = -(3x-2y)^2 + \left(\sqrt{29}y\right)^2.$$

অবশ্য কাজটা অন্যভাবেও করা যেত। যেমন ধরো, যদি y থেকে শুরু করতাম--

$$-9x^2 + 12xy + 25y^2 = \left(5y + \tfrac{6x}{5}\right)^2 - \left(\tfrac{\sqrt{261}}{5}x\right)^2.$$

কিন্তু যেভাবেই এগোও, কাজটা করা যাবেই। তবে যদি অনেকগুলো variable থাকে, তবে মাথা গুলিয়ে যাবার সম্ভাবনা। তখন matrix দিয়ে ভাবলে সুবিধা হবে। এই কায়দাটাই এবার আমরা শিখব। এই উদাহরণটাই আরেকবার দেখি।

$$-9x^2 + 12xy + 25y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} -9 & 6 \\ 6 & 25 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

আমরা কাজ করব মাঝের matrix-টা নিয়ে। আমরা পুরো জিনিসটাকে খালি square দিয়ে লিখতে চাই, কোন গুণফল জাতীয় জিনিস (যেমন $12xy$) চাই না। সুতরাং এখানে আমাদের উদ্দেশ্য হবে off-diagonal entry-গুলোকে শূন্য করে দেওয়া। তার জন্য আমরা প্রথম row-এর $\frac{2}{3}$ যোগ করে দেব দ্বিতীয় row-এর সঙ্গে--

$$\begin{bmatrix} -9 & 6 \\ 0 & 29 \end{bmatrix}.$$

এরপর একই কাজ করব column-গুলোর বেলাতেও, মানে প্রথম column-এর $\frac{2}{3}$ যোগ করে দেব দ্বিতীয় column-এ--

$$\begin{bmatrix} -9 & 0 \\ 0 & 29 \end{bmatrix}.$$

এর মানে হল আমরা গোড়ার quadratic form-টাকে এইভাবে লিখছি--

$$-9x^2 + 12xy + 25y^2 = -9x_1^2 + 29y_1^2,$$

যেখানে x_1 আর y_1 হল x, y-কে মিলিয়ে মিশিয়ে তৈরী। যেমন এক্ষেত্রে হিসেব করে দেখলে পাবে $x_1 = x - \frac{2}{3}y$ আর $y_1 = y$. এইভাবে x_1, y_1-কে x, y দিয়ে লেখাটা সব সময়ে সহজ নাও হতে পারে, কিন্তু অনেক সময়েই সেটা গুরুত্বপূর্ণ নয়, গুরুত্বপূর্ণ হল quadratic form-টার মূল কাঠামোটা--সেটা কি কয়েকটা square-এর যোগফল, নাকি বিয়োগফল। যেমন নীচের তিনটে integral-এর দিকে একবার তাকিয়েই তুমি বলে দিতে পারো কোথায় integration-এর কোন্ ফর্মুলাটা লাগবে--

$$\int \frac{dx}{\sqrt{(x+1)^2 - 34}}, \int \frac{dx}{\sqrt{(2x+19)^2 + 3}}, \int \frac{dx}{\sqrt{-(2x+3)^2 + 25}}.$$

আরেকটা উদাহরণ হল ellipse আর hyperbola চেনা। তুমি জানো যে, এইরকম equation-গুলো হয় ellipse-এর--

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

আর এরা হয় hyperbola–

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

সুতরাং ellipse আর hyperbola-কে আলাদা করার জন্য খালি square দুটোর চিহ্ন জানাই যথেষ্ট। আমাদের উদাহরণেও যদি তুমি খালি চিহ্নটুকুই চাও তবে $-9x_1^2 + 29y_1^2$-কে লিখতে পারি $-x_2^2 + y_2^2$, যেখানে $x_2 = 3x_1$ এবং $y_2 = \sqrt{29}y_1$. মানে 9 আর 29-কে আমরা variable-গুলোর মধ্যে শুষে নিলাম, খালি চিহ্নটুকুই বাইরে রইল। লক্ষ্য কর 9-কে শুষে নেওয়ার জন্য square-এর মধ্যে x_1-কে আমরা $\sqrt{9} = 3$ দিয়ে গুণ করলাম। একই কাজ আমরা matrix দিয়েও করতে পারতাম। আমরা পেয়েছিলাম

$$\begin{bmatrix} -9 & 0 \\ 0 & 29 \end{bmatrix},$$

এর প্রথম row-কে তুমি $\sqrt{9} = 3$ দিয়ে ভাগ করে দাও, তারপর একই কাজ করো প্রথম column-এর উপরে। ফলে -9 সংখ্যাটা দুবার 3 দিয়ে ভাগ হয়ে -1 হয়ে যাবে। একইভাবে দ্বিতীয় row এবং column-কে ভাগ কর $\sqrt{29}$ দিয়ে। তবে পাবে

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

অনেক সময়ে লোকের এতেও শান্তি হয় না, আরও গুছিয়ে লিখতে চায়। যেমন প্রথমে 1-গুলো, তারপরে -1-গুলো লিখতে ভালোবাসে। তার জন্য row-দুটোকে interchange (আগে-পরে) করে দাও, এবং column-গুলোকেও তাই কর। এইবার পাবে--

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

এটাকে আমরা বলব মূল matrix-টার **normal form** বা আরো গুছিয়ে বললে, normal form under congruence. আমরা আগে বলেছিলাম যে square complete করার সময়ে অনেকভাবে এগোনো যায়। বিভিন্নভাবে এগোলে বিভিন্নরকম উত্তর আসে। কিন্তু normal form-এর ক্ষেত্রে সে সমস্যা নেই, যেভাবেই এগোও শেষ পর্যন্ত normal form-টা একই আসবে। আমরা অংকের ভাষায় গুছিয়ে normal form-এর সংজ্ঞা এক্ষুণি লিখব। তার আগে আরেকটা উদাহরণ দেখে নিই। এবার মনে কর quadratic form-টা হল $-y^2 - (z-x)^2$. এখানে তো square করেই দেওয়া আছে। দেখাই যাচ্ছে যে তিনটে variable, আর দুটো square, দুজনেরই চিহ্ন '$-$'। তার মানে এখানে normal form হবে

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

56.2.2 অংকের ভাষায়

আমরা দেখলাম যে, normal form হল একটা diagonal matrix, যার diagonal-এ $1, -1, 0$ ছাড়া আর কিছু নেই। যদি কোনো 1 থাকে, তবে তারা থাকবে একেবারে শুরুতে। কোনো -1 থাকলে তারা আসবে এর পরে, আর 0 থাকলে সবার শেষে আসবে। এই কথাটা গুছিয়ে লিখে দিলেই normal form-এর সংজ্ঞা পেয়ে যাব--

> DEFINITION: **Normal form (under congruence)**
>
> A symmetric matrix is said to be in **normal form** (under congruence) if it is a diagonal matrix with nothing but $1, -1$ and 0 in the diagonal entries. The 1's (if any) will come first, then the -1's (if any), then the 0's (if any).

এখানে ব্র্যাকেটের মধ্যে "under congruence" লিখেছি, কারণ যেকোনো equivalence-র জন্য normal form-এর আলোচনা করা যায়। তাই খালি normal form বললে প্রশ্ন এসে পড়ে equivalence-টা কী। অবশ্য এই অধ্যায়ে আমরা অন্য কোনো normal form-এর আলোচনা করব না, তাই "under congruence" কথাটা এর পর থেকে উহ্য রাখব।
আমরা আগেই জানি যে, কোনো diagonal matrix-এর কতগুলো nonzero diagonal entry রয়েছে, সেটাই হল তার rank. সুতরাং এখানে 1 এবং -1-এর মোট সংখ্যাই হবে rank.

> THEOREM
>
> The rank of a matrix in normal form is the number of nonzero diagonal elements.

কতগুলো 1 আছে, সেই সংখ্যাটাকে আলাদা করে বলে **index**, এবং তা থেকে থেকে -1-এর সংখ্যা বিয়োগ করলে যা থাকে সেটাকে বলে **signature**.

> DEFINITION: **Index, signature**
>
> For a matrix in normal form we define
>
> 1. the **index** as the the number of 1's in the diagonal,
> 2. the **signature** as the the number of 1's in the diagonal minus the number of -1's in the diagonal.

যেমন একটু আগেই যে normal form-টা দিলাম, তার index হল 0 (কারণ কোনো 1 নেই diagonal-এ), rank হল $0 + 2 = 2$, আর signature হল $0 - 2 = -2$.

Exercise 10: একটা diagonal matrix-এর diagonal entry-গুলো হল $2, 3, -4, 0, 4, -5, 0$. এর rank, index এবং signature বার করো। ■

56.3 Uniqueness

আমরা জানি যে, কোনো ভগ্নাংশের ঠিক একটাই reduced form হয়। এই uniqueness-টাকে ঘুরিয়ে এভাবে বলা যায়--

> যদি $\frac{p}{q}$ এবং $\frac{m}{n}$ দুজনেই একই ভগ্নাংশ বোঝায়, এবং দুজনেই reduced form-এ থাকে, তবে $p = m$ এবং $q = n$ হতে বাধ্য।

তেমনই, কোনো symmetric matrix-এর ঠিক একটাই normal form থাকে। এই uniqueness-টাকে ঘুরিয়ে লিখতে পারি--

> যদি A, B দুজনেই symmetric matrix হয়, যারা দুজনেই normal form-এ আছে এবং পরস্পরের সঙ্গে congruent, তবে $A = B$ হতে বাধ্য।

এখানে A আর B দুজনেই normal form-এর থাকার ফলে $A = B$ দেখানো মানে এটা দেখানো যে, ওদের দুজনের diagonal-এই সমান সংখ্যক 1, সমান সংখ্যক -1 এবং সমান সংখ্যক 0 আছে। এদের মধ্যে সমান সংখ্যক 0 দেখানোটা বেশ সহজ। সেই প্রমাণটা নিয়েই নীচের অংকটা।

Exercise 11: ধরো A, B দুজনেই normal form-এ আছে এবং ওরা পরস্পরের সঙ্গে congruent. বুঝতেই পারছ A, B দুজনেরই সাইজ সমান।

1. দেখাও যে $r(A) = r(B)$.
2. এ থেকে সিদ্ধান্ত কর যে, A আর B-র diagonal বরাবর 0-র সংখ্যাও সমান হতে বাধ্য।

HINT:
কোনো nonsingular matrix দিয়ে গুণ করলে একটা matrix-এর rank-এর কীরকম পরিবর্তন হয়?
■

অবশ্য এতে এটা প্রমাণ হচ্ছে না যে, diagonal বরাবর 1-এর সংখ্যা এবং -1-এর সংখ্যা আলাদা করে সমান। সেটা আসে নীচের theorem-টা থেকে--

Sylvester's law of inertia

Let A, B be real symmetric matrices congruent to each other. Then the number of positive (negative, zero) eigenvalues of A must equal the number of positive (negative, zero) eigenvalues of B.

এটা থেকে normal form-এর uniqueness যে অতি সহজেই আসছে, সেটা আশা করি বুঝতে পারছ। কারণ যদি A আর B দুজনেই normal form-এ থাকে, তবে ওদের diagonal entry-গুলোই ওদের eigenvalue. সুতরাং positive eigenvalue-র সংখ্যা মানে diagonal বরাবর 1-এর সংখ্যা, ইত্যাদি।
Theorem-টার প্রমাণ খুব সহজ নয়, তাই আমরা ও পথে যাব না।
যে Sylvester-এর নাম এই theorem-এর সঙ্গে জড়িত, সেই ভদ্রলোকের আরেকটি অবদান আমরা না জেনেও মুহূর্মুহূ ব্যবহার করছি। এই যে আমরা এত matrix নিয়ে কাজ করছি, সেই "matrix" নামটা কিন্তু প্রথম Sylvester-ই দিয়েছিলেন।

DAY 57
Congruence (part 2)

57.1 ধাপে ধাপে normal form-এ পৌঁছনো

ভগ্নাংশদের বেলায় যেমন কাটাকাটি করে যেকোনো ভগ্নাংশকেই reduced form-এ পরিণত করা যেত, তেমনি যেকোনো symmetric matrix-কেই ধাপে ধাপে normal form-এ পরিণত করার একটা কায়দা আছে। কায়দাটা ভিতরে ভিতরে সেই square complete করাই। কিন্তু matrix ব্যবহার করে করা হয় বলে Gauss-Jordan elimination-এর মত elementary operation-গুলো ব্যবহৃত হয়। একটা উদাহরণ দিয়ে বোঝাই।

Example 10: Reduce the real quadratic form $2x^2 + 2y^2 + 5z^2 - 4xy - 2xz + 2yz$ to its normal form and find its rank, signature.

SOLUTION: সবার প্রথম কাজ হল quadratic form-টাকে matrix দিয়ে লেখা--

> The quadratic form is
>
> $$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & -2 & -1 \\ -2 & 2 & 1 \\ -1 & 1 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

এবার আমাদের কাজ এগোবে তিন পর্যায়ে। প্রথম পর্যায়ে আমরা matrix-টাকে diagonal বানাব। দ্বিতীয় পর্যায়ে diagonal entry-গুলোকে 1 বা -1 বা 0 বানাব। তৃতীয় পর্যায়ে সেই diagonal entry-দের সুন্দর করে সাজাব, শুরুতে 1-গুলো, তারপর -1-গুলো, এবং শেষে 0-গুলো। প্রথম পর্যায়টা শুরু করা যাক। এর মধ্যেও কয়েকটা ধাপ আছে, matrix-টার যদি $n \times n$ হয়, তবে ধাপের সংখ্যা হবে n.

> We apply congruence operations to the matrix to reduce it to diagonal form.
>
> $$\begin{bmatrix} 2 & -2 & -1 \\ -2 & 2 & 1 \\ -1 & 1 & 5 \end{bmatrix}$$
>
> $$\xrightarrow[]{\substack{newR_2 = R_2 + R_1 \\ newR_3 = R_3 + \frac{1}{2}R_1}} \begin{bmatrix} 2 & -2 & -1 \\ 0 & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & \mathbf{\frac{9}{2}} \end{bmatrix}$$

এইটা কী করলাম একটু বুঝে দেখি। আমরা matrix-টাকে diagonal বানানোর চেষ্টা করছি, তার মানে off-diagonal সংখ্যাগুলোকে শূন্য বানাতে চাই। প্রথম ধাপে আমরা প্রথম column-এর নীচের সংখ্যাদুটোকে (মানে -2 আর -1-কে) শূন্য বানাচ্ছি। তার জন্য প্রথম row-টাকে দ্বিতীয় row-র সাথে যোগ করেছি (যাতে -2 আর 2 কেটে যায়)। এই কাজটাকেই সংক্ষেপে লিখেছি--

$$newR_2 = R_2 + R_1.$$

এতে -2-টা বিদায় হল। -1-টাকে তাড়াবার জন্য প্রথম row-এর অর্ধেক যোগ করেছি তৃতীয় row-তে।

ভালো করে লক্ষ্য কর যে, 3×3 matrix-এর 9-টা সংখ্যার মধ্যে খালি চারটেকে কষ্ট করে বার করতে হল। খেয়াল রাখার সুবিধার জন্য এগুলোকে মোটা করে দেখিয়েছি। এখানে প্রথম row-টা অপরিবর্তিত, আর বাকী row-দুটোর প্রথম সংখ্যা 0 হতে বাধ্য।

এবার congruent column operation করব, অর্থাৎ যা যা row-এর উপর করলাম, ঠিক তাই তাই column-এর উপর করব।

$$\xrightarrow[\;]{\substack{newC_2 = C_2 + C_1 \\ newC_3 = C_3 + \frac{1}{2}C_1}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{9}{2} \end{bmatrix}$$

Congruent column operation করার জন্য কোনোই পরিশ্রম নেই, খালি প্রথম row-এর শেষ দুটো সংখ্যাকে 0 করে দেওয়া। এই ধাপে আমাদের মূল হাতিয়ার ছিল (1,1)-th entry-টা। ওটাকে ব্যবহার করেই আমরা ওর নীচের আর ডানদিকের entry-গুলোকে শূন্য বানালাম। এই (1,1)-th entry-টাকে বলব আমাদের প্রথম ধাপের **pivot.** যে row-তে (বা column-এ) pivot-টা আছে, তাকে অনেক সময়ে pivotal row (বা column) বলে। প্রতিধাপে pivot-এর অবস্থান বদলায়। যখন আমরা i-th ধাপে থাকি, তখন pivot হয় (i,i)-th entry-টা।

প্রতি ধাপে আমাদের লক্ষ্য থাকে pivot-এর নীচের আর ডানদিকের entry-গুলোকে শূন্য বানানো। আমাদের প্রথম ধাপ শেষ হয়ে গেছে, কারণ এখন (1,1)-th entry-র নীচে আর ডানদিকে সবাই শূন্য। এবার দ্বিতীয় ধাপ, মানে $i = 2$. এখন pivot হল (2,2)-th entry. এর নীচে আর ডানদিকে তো সবাই শূন্য হয়েই আছে, তাই এই ধাপে করার কিছু নেই। সবার শেষ ধাপে $i = 3$. শেষ ধাপে কখনোই কিছু করার থাকে না, কারণ এই ধাপে pivot-এর নীচে বা ডানদিকে কোনো entry-ই নেই।

আমাদের matrix-টা diagonal হয়ে গেছে, অতএব প্রথম পর্যায় শেষ। এবার দ্বিতীয় পর্যায়, মানে positive diagonal element-দের 1 আর negative-দের -1 বানানো। এখানে কোনো negative কিছু নেই, সুতরাং খালি positive-দের 1 বানালেই চলবে। এর জন্য কায়দা হল যে সংখ্যাটাকে 1 বানাব, তার row আর column-কে সংখ্যাটার square root দিয়ে ভাগ করে দেওয়া। দুবার square root দিয়ে ভাগ করলে সংখ্যাটা কেটে গিয়ে 1 হয়ে যেতে বাধ্য!

$$\xrightarrow[\;]{\substack{newR_1 = R_1/\sqrt{2} \\ newR_3 = R_3/\sqrt{\frac{9}{2}} \\ newC_1 = C_1/\sqrt{2} \\ newC_3 = C_3/\sqrt{\frac{9}{2}}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

এখানে আমাদের মোট চারটে operation লেগেছে, দুটো row operation আর দুটো column operation. কিন্তু দেখতেই খালি চারটে operation, আসলে কোনো অংকই কষতে হয় নি, খালি positive diagonal-গুলোর জায়গায় 1 বসিয়েই হয়ে গেছে (negative diagonal element থাকলে তাদের জায়গায় -1 বসাতাম)।

ব্যস্ , দ্বিতীয় পর্যায় শেষ। এই পর্যায়টা সবসময়েই এক লাইনে করা যায়।

সর্বশেষ পর্যায় হল diagonal element-গুলোকে সাজিয়ে লেখা, যাতে প্রথমে 1-রা তার পরে -1-রা থাকে (এখানে অবশ্য -1-এর ব্যাপার নেই)। এর জন্য দরকার খালি কিছু row-এর (এবং symmetry রক্ষা করতে column-দেরও) স্থানবিনিময় করা।

$$\xrightarrow[\;]{\substack{newR_2 = R_3 \\ newR_3 = R_2 \\ newC_2 = C_3 \\ newC_3 = C_2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

এবার যা যা বার করতে বলেছে, সেগুলো বার করতে আর কোনো অসুবিধা নেই--

> Rank = no. of nonzero diagonals = 2.
>
> Signature = no. of positive diagonals - no. of negative diagonals = $2-0=2$.

■

এই অংকে যে কায়দাটা করলাম, তার মূলমন্ত্র বলা যেতে পারে--pivot-দের ব্যবহার করে তাদের নীচের এবং ডানদিকের entry-দের শূন্য বানানো। এই কাজে খালি একটাই সমস্যা হতে পারে, যেটা আমাদের উদাহরণটায় হয়নি। সেটা হল যদি কখনো কোনো pivot শূন্য হয়ে যায়, অথচ তার নীচে এবং ডানদিকে কোনো non-zero entry থাকে। সেরকম ক্ষেত্রে কী করতে হবে, সেটা একটা উদাহরণ দিয়ে দেখাই।

Example 11: ধরো matrix-টা হল $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. এটার normal form বার করো।

SOLUTION: এখানে প্রথম ধাপেই সেই সমস্যা। $(1,1)$-th entry-টা শূন্য কিন্তু (1,2)-th আর (2,1)-th entry হল $1 \neq 0$. এখানে আমরা ওই non-zero off-diagonal entry-র row এবং (symmetry বজায় রাখতে column) যোগ করে দেব pivotal row এবং column-এর সাথে। মানে এখানে দ্বিতীয় row যোগ করব প্রথম row-এর সাথে। ফলে পাব $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. এবার একই কাজ করব column-দের নিয়েও। ফলে আসবে $\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$. ব্যস্, এবার diagonal entry-টা non-zero হয়ে গেছে, তাই আগের অংকের মতই এগোতে পারবে। ■

Exercise 12: আচ্ছা, এই যে অংকটা দেখলে, সেখানে কি প্রথম row-এর সঙ্গে দ্বিতীয় row যোগ না করে, ওই দুটো row-র স্থানবিনিময় করে দিলেই চলত না? ■

একই রকম আরো কয়েকটা অংক। তোমার হাত পাকানোর জন্য রেখে দিলাম।

Exercise 13: Find the normal form and hence the rank and signature of the real quadratic form $3x^2 + 8y^2 - 24z^2 + 10yz + 10zx - 14xy$. ■

Exercise 14: Find the canonical form of the quadratic form $xy - 3yz + zx$. Find also its rank, index and signature. ■

Exercise 15: Reduce the equation $2x^2 + 5y^2 + 10z^2 + 4xy + 6xz + 12yz$ into its canonical form.[3] **(2021.CC12.3ci)** ■

Exercise 16: Reduce the quadratic $Q = xy + yz + zx$ to its canonical form and hence find its rank and signature.

■

Exercise 17: Obtain the normal form under congruence of the following matrix and find its rank, index and signature:

$$\begin{bmatrix} 4 & 5 & -1 & 0 \\ 5 & 0 & 2 & 1 \\ -1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 1 \end{bmatrix}$$

■

Exercise 18: Choose the correct answer with proper justification (1 mark for right answer and 1 mark for justification). Which of the following is the signature of the quadratic form $xy + yz + zx$?

(A) 1 (B) -1 (C) 2 (D) -2.

(2022.CC12.1h) ■

Exercise 19: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). Signature of the quadratic form $5x^2 + y^2 + 10z^2 - 4yz - 10zx$ is

(A) 1 (B) 2 (C) 3 (D) 4.

(2020.CC12.1i) ■

57.2 Transformation matrix-টা বার করা

আমরা দেখেছি যে, কোনো symmetric matrix দেওয়া থাকলে আমরা তার জন্য normal form বার করতে পারি। যদি matrix-টাকে A বলি, তবে $P'AP = D$ লিখতে পারি, যেখানে P একটা nonsingular matrix এবং D আছে normal form-এ। এতক্ষণ আমরা D বার করার কায়দাকানুন নিয়ে আলোচনা করছিলাম। এবার সেই সঙ্গে P-ও বার করতে চাই। এতক্ষণ যেভাবে কোনো real, symmetric matrix-কে normal form-এ পরিণত করছিলাম, সেখানে elementary row বা column operation ব্যবহার করা হচ্ছিল। আমরা জানি যে row operation-দের বাঁদিক থেকে matrix দিয়ে গুণ বলে ভাবা যায়, এবং column operation-দের বেলায় ডানদিক থেকে। সুতরাং A যদি একটা real, symmetric matrix হয়, যার normal form হল D, তবে

$$QAP = D,$$

যেখানে Q হল row operation-দের matrix আর P হল column operation-দের। যেহেতু operation-গুলো congruent-ভাবে করা হয়েছিল, তাই এক্ষেত্রে $Q = P'$ হবে, মানে

$$P'AP = D.$$

এই P-টা বার করা খুবই সহজ, ঠিক যেভাবে আমরা inverse বার করতাম, A-র গায় একটা I জুড়ে দিয়ে।

Example 12: Find the associated matrix A of the quadratic form

$$y^2 + z^2 + 2xy + 2yz + 4zx.$$

Find a matrix P such that P^TAP is the normal form of A.

SOLUTION:

$$y^2+z^2+2xy+2yz+4zx = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Thus the associated matrix is

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

এটাও আগের অংকগুলোর মতই, খালি এবার বাড়তির মধ্যে P-টা চেয়েছে। P^TAP-তে P^T আছে A-র বাঁদিকে তার মানে P^T matrix-টা করছে row operation-গুলো। যদি আমরা row operation-গুলোর হিসেব রাখি, তাহলেই P^T পেয়ে যাব। সেই কারণে এবার আমরা A-র ডানদিকে একটা identity matrix জুড়ে নেব।
গোড়াতেই সমস্যা, pivot-টা শূন্য হয়ে গেছে। তার নীচের যা খুশি একটা non-zero entry নেব (এখানে (2,1)-th entry-টা নিয়েছি)। তার row-টাকে pivotal row-র সাথে যোগ করে দেব।

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{newR_1 = R_2 + R_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

অবশ্যই আমাদের congruent column operation-টাও করতে হবে, নইলে matrix-টা symmetric থাকবে না। লক্ষ্য কর যে, matrix-টার ডানদিকের অংশটার উপর column operation-এর কোনো প্রভাব নেই। ওই অংশটা আছে খালি row operation-গুলোর হিসেব রাখার জন্য। Row operation-দের হিসেব থাকলে column operation-গুলোর হিসেব এমনিতেই বেরিয়ে যাবে, কারণ column operation-গুলো হচ্ছে row operation-গুলোর congruent.

$$\xrightarrow{newC_1 = C_2 + C_1} \left[\begin{array}{ccc|ccc} 3 & 2 & 3 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

এবার যেমন আগের অংকগুলো করেছিলাম সে ভাবেই এগোতে পারব। এখানে বলে রাখি যে প্রথম row-র প্রথম সংখ্যাটাকে nonzero বানাবার আরো কায়দা আছে, যেমন এক্ষেত্রে প্রথম row-দুটোকে swap করে দিলেও হত (এবং অবশ্যই প্রথম column দুটোকেও congruent-ভাবে swap করতে হবে)।

$$\xrightarrow[\substack{newR_2 = R_2 - \frac{2R_1}{3} \\ newR_3 = R_3 - R_1}]{} \left[\begin{array}{ccc|ccc} 3 & 2 & 3 & 1 & 1 & 0 \\ 0 & -\frac{1}{3} & -1 & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & -1 & -2 & -1 & -1 & 1 \end{array}\right]$$

$$\xrightarrow[\substack{newC_2 = C_2 - \frac{2C_1}{3} \\ newC_3 = C_3 - C_1}]{} \left[\begin{array}{ccc|ccc} 3 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{3} & -1 & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & -1 & -2 & -1 & -1 & 1 \end{array}\right]$$

$$\xrightarrow{newR_3 = R_3 - 3R_2} \left[\begin{array}{ccc|ccc} 3 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{3} & -1 & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{newC_3 = C_3 - 3C_2} \left[\begin{array}{ccc|ccc} 3 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{3} & 0 & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right]$$

$$\xrightarrow[\substack{\text{Swap } R_2, R_3 \\ \text{Swap } C_2, C_3}]{} \left[\begin{array}{ccc|ccc} 3 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & -2 & 1 \\ 0 & 0 & -\frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \end{array}\right]$$

$$\xrightarrow[\substack{newR_1 = R_1/\sqrt{3} \\ newR_3 = \sqrt{3}R_3 \\ newC_1 = C_1/\sqrt{3} \\ newC_3 = \sqrt{3}C_3}]{} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 & 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & -\frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 \end{array}\right]$$

Thus, the required normal form is

$$P^T AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

and

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & 1 & -\frac{2}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & -2 & \frac{1}{\sqrt{3}} \\ 0 & 1 & 0 \end{bmatrix}.$$

এই P-টা পেলাম vertical লাইনের ডানদিকে যে matrix-টা এসেছিল (মানে P^T), তাকে transpose করে। ■

Exercise 20: Find the the matrix P such that the congruent matrix $B = P^t AP$ is a diagonal matrix if

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & -3 \end{bmatrix}.$$

[P^t is the transpose of the matrix P.]

HINT:
ঠিক আগের অংকটার মত এগোলেই হবে। তবে এখানে normal form অব্ধি যেতে বলেনি, খালি diagonal পর্যন্ত গেলেই চলবে। তার মানে আগের অংকের শেষ ধাপদুটো না করলেও চলবে। ■

Exercise 21: Diagonalise the symmetric matrix $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$.[5] **(2020.CC12.3f)**

■

Exercise 22: Diagonalise the matrix $\begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 3 \end{bmatrix}$.[5] **(2022.CC12.3g)** ■

DAY 58
বিভিন্ন ধরণের quadratic form

ধরো এই quadratic form-টা তোমায় দিলাম--

$$x^2 + 2y^2 + 5z^2.$$

এমন $x, y, z \in \mathbb{R}$ দিতে পারো, যাতে এটা < 0 হয়? অসম্ভব, কারণ এটা সবচেয়ে কম হতে পারে 0, তাও যদি x, y, z সবাই একসঙ্গে 0 হয়। তার মানে যদি x, y, z-এর একজনও nonzero হয়, তবে তুমি নিশ্চিত হয়ে বলতে পারো যে, এটা positive হবেই। "নিশ্চিত" হওয়াকে ইংরাজিতে বলে "definite" হওয়া। তাই এই quadratic form-টাকে আমরা বলব **positive definite.** কিন্তু যদি এই quadratic form-টা দিতাম--

$$x^2 + (y - z)^2,$$

তবে কিন্তু তুমি অত নিশ্চিত হতে পারতে না, কারণ যদি $x = 0$ আর $y = z = 1$ নিই, তবে এটা 0 হয়ে যাবে। অবশ্য এটা কোনোভাবেই < 0 হতে পারে না। সুতরাং জিনিসটা অন্ততঃ nonnegative যে হবে সেটুকু নিশ্চিত। তাই একে আমরা বলব **nonnegative definite.** অবশ্য আমরা জানি যে, একটা সংখ্যা যদি positive হয়, তবে সেটা nonnegative-ও হবে। সুতরাং কোনো quadratic form যদি positive definite হয়, তবে সেটা nonnegative definite-ও হবে। যদি কোনো quadratic form দেখো যেটা nonnegative definite কিন্তু positive definite নয় (যেমন এক্ষুণি যে উদাহরণটা দেখলাম), তাদের অনেক সময়ে লোকে বলে **positive semidefinite,** যদিও এটা খুব বেশী প্রচলিত ভাষা নয়। আমরা positive আর nonnegative-দের নিয়ে যা যা করলাম, সে সবই negative আর nonpositive-দের নিয়েও করতে পারি। যেমন $-x^2 - y^2 - z^2$ হল x, y, z-এর এমন একটা quadratic form যেটা **negative definite.** সংজ্ঞাগুলো চটপট লিখে ফেলি।

58.1 সংজ্ঞা

অনেকগুলো সংজ্ঞা হবে, তিনভাগে ভেঙে লিখি--

Definition: Definiteness (part 1)

A real, symmetric matrix A is called

- **positive definite (pd)** if
$$\forall \vec{x} \neq \vec{0} \quad \vec{x}' A \vec{x} > 0.$$
- **nonnegative definite (nnd)** if
$$\forall \vec{x} \neq \vec{0} \quad \vec{x}' A \vec{x} \geq 0.$$
- **positive semidefinite (psd)** if it is nnd but not pd.

এখানে nnd-র সংজ্ঞায় $\forall \vec{x} \quad \vec{x}' A \vec{x} \geq 0$ লিখলেও চলত, কারণ $\vec{x} = \vec{0}$ হলে $\vec{x}' A \vec{x} = 0$ হবে। এই গেল বিভিন্ন রকম positive আর nonnegative-দের গল্প। এবার negative-দের কথা শোনা যাক।

Definition: Definiteness (part 2)

A real, symmetric matrix A is called

- **negative definite (nd)** if
$$\forall \vec{x} \neq \vec{0} \quad \vec{x}' A \vec{x} < 0.$$
- **nonpositive definite (npd)** if
$$\forall \vec{x} \neq \vec{0} \quad \vec{x}' A \vec{x} \leq 0.$$
- **negative semidefinite (nsd)** if it is npd but not nd.

যারা positive, negative কোনোদিকেই নেই, তাদের কথা রয়েছে নীচের সংজ্ঞাটায়--

Definition: Definiteness (part 3)

A real, symmetric matrix A is called **indefinite (id)** if it is neither nnd nor npd.

এই তো গেল সংজ্ঞা। বুঝতেই পারছ যে, যেকোনো real, symmetric matrix-ই এই সাত রকমের মধ্যে পড়বে। ছবি এঁকে ভাবলে ব্যাপারটা Fig 23-এর মত।

Exercise 23: Fig 23-এ লক্ষ্য কর, যাবতীয় nonnegative definite matrix আর যাবতীয় nonpositive definite matrix-এর set-এর মধ্যে একটা intersection রয়েছে। মানে এমন real, symmetric matrix সম্ভব যেটা একাধারে nonnegative definite এবং nonpositive definite. আসলে এরকম ঠিক একটাই matrix আছে। বলতে

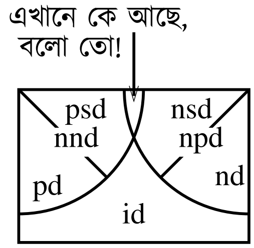


Fig 23

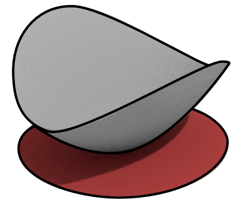


Fig 24

পারো সেটা কী? ■

58.2 ছবি দিয়ে বোঝা

এবার পুরো জিনিসটাকে গ্রাফ দিয়ে এঁকে আলোচনা করি। মনে কর একটা quadratic form দিলাম--

$$f(x,y) = ax^2 + bxy + cy^2.$$

এটা x, y-এর একটা function. এর গ্রাফ একটা surface-এর মত হবে। এই surface-টা দেখলেই অনায়াসে বলে দেওয়া যায় quadratic form-টা কোন্ ধরণের। যেমন Fig 24-এর কথাই ধরো। এখানে চাকতির মত জিনিসটা হল xy-plane-এর একটা গোলাকার অংশ, যার কেন্দ্রে আছে $(0,0)$. আর তার উপরে বাটির মত জিনিসটাই হল আমাদের গ্রাফটা। লক্ষ্য কর, বাটিটা সম্পূর্ণই চাকতির উপরে আছে, এবং খালি $(0,0)$-এ চাকতিকে স্পর্শ করেছে। মানে সর্বত্রই $f(x,y) > 0$, একমাত্র ব্যতিক্রম $f(0,0) = 0$. তাই এখানে quadratic form-টা হল positive definite. যদি Fig 25-এর দিকে তাকাও, তবে দেখবে গ্রাফটা আর বাটির মত নেই, কতকটা ধোসার মত হয়ে গেছে। এবং নীচের দিকটা একটা পুরো লাইন বরাবর প্লেটের সঙ্গে (মানে xy-plane-এর সঙ্গে) লেপ্টে আছে। অর্থাৎ সর্বত্রই $f(x,y) \geq 0$, এবং এমন $(x,y) \neq (0,0)$ আছে, যাতে $f(x,y) = 0$ হয়। ফলে এখানে quadratic form-টা হল positive semidefinite. যদি Fig 24-র বাটিটাকেই উল্টে চাকতিটার তলায় নিয়ে যাও (Fig 26-র মত), তবে যে একটা negative definite জিনিস পাবে, সেটা নিশ্চয়ই আর বলে দিতে হবে না? একইভাবে Fig 25-কে ওল্টালে negative semidefinite একটা উদাহরণ পাবে, সেটা আর আলাদা করে আঁকলাম না। এবার Fig 27-এর দিকে তাকাও। এখানে লক্ষ্য কর যে, গ্রাফটার কিছু অংশ xy-plane-এর উপরে, কিছু অংশ নীচে। মানে এখানে $f(x,y) > 0$-ও হচ্ছে আবার $f(x,y) < 0$-ও হচ্ছে। তাই এটা indefinite একটা quadratic form-এর উদাহরণ।

58.3 Normal form ব্যবহার করে বার করা

এবার ধরো তোমাকে একটা real, symmetric matrix ধরিয়ে দিয়ে জানতে চাইলাম সেটা এই সাতটা দলের মধ্যে কোথায় পড়ে। কী করে এগোবে? একটা সহজ কায়দা হল normal form-টা বার করে ফেলা। যদি এর normal form-এর diagonal-এ--

1. খালি 1 থাকে (কোনো -1 বা 0 না থাকে), তবে (এবং একমাত্র তবেই) positive definite হবে। অর্থাৎ pd-দের

Fig 25

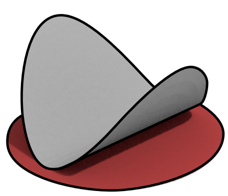


Fig 26

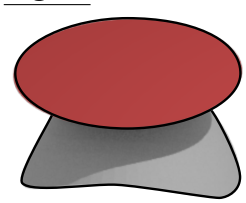
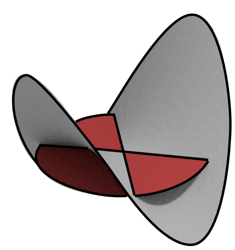


Fig 27



বেলায় normal form সর্বদা identity matrix হয়।

2. যদি কোনো -1 না থাকে, তবে (এবং একমাত্র তবেই) nonnegative definite হবে।

3. একইভাবে যদি খালি -1 থাকে, তবে (এবং একমাত্র তবেই) negative definite হবে, ইত্যাদি।

4. আর যদি দেখো যে, 1 এবং -1 দুইই আছে, তবে সেই matrix-এর পক্ষে indefinite হওয়া ছাড়া পথ নেই। বিপরীতপক্ষে, যেকোনো id matrix-এর normal form-এর diagonal-এই অন্ততঃ একটা 1 এবং অন্ততঃ একটা -1 থাকবেই।

কয়েকটা উদাহরণ দেখি।

Exercise 24: Reduce the quadratic form $6x^2 + y^2 + 18z^2 - 4yz - 12zx$ to normal form and show that it is positive definite.
HINT:
একটা quadratic form দিয়ে দেখাতে বলা হয়েছে যে সেটা positive definite. এর জন্য আমরা প্রথমে quadratic form-টার associated matrix-টা লিখব, এবং সেটাকে ধাপে ধাপে normal form-এ পরিণত করব। যদি normal form-এর সব গুলো diagonal entry-ই 1 হয় (মানে normal form-টা identity matrix হয়), তবেই quadratic form-টা positive definite হবে।

$$\begin{bmatrix} 6 & 0 & -6 \\ 0 & 1 & -2 \\ -6 & -2 & 18 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

এখানে normal form-এ পৌঁছনোর ধাপগুলো আর দেখালাম না। না বুঝলে উত্তর দেখে নিও। যেহেতু normal form-টা identity matrix এল, তাই মূল matrix-টা pd হবে। ■

Exercise 25: Reduce the quadratic form $xy + 2yz + 2zx$ to normal form. Find its rank and nature. (x, y, z are real numbers.)
HINT: এখানে nature বার করা মানে হল সেই সাতটা দলের কোন্ দলে পড়বে, সেটা বার করা। ■

58.4 P.d. আর n.n.d.

আমরা যে সাত রকমের definite-এর গল্প শিখলাম, তাদের মধ্যে indefinite-রা সাধারণতঃ সবচেয়ে বিচ্ছিরি। ওদের নিয়ে কাজ করা কঠিন। Positive definite আর negative definite-দের আচরণ একইরকম, বস্তুতঃ তুমি যা খুশি একটা negative definite matrix নিয়ে, তার আগে একটা মাইনাস বসিয়ে দাও, অমনি সেটা একটা positive definite matrix হয়ে যাবে। সেই কারণে লোকে negative definite-দের নিয়ে আর আলাদা করে মাথা ঘামায় না। একই কারণে nonnegative definite-দের নিয়ে মাথা ঘামালেই nonpositive definite-দেরও হিল্লে হয়ে যায়। আর যদি nonnegative definite এবং positive definite-দের সঙ্গে পরিচয় থাকে, তবে positive semidefinite-দের চিনতেও কোনো কষ্ট নেই। সেই কারণে লোকে খালি positive definite আর nonnegative definite-দের নিয়েই বেশী পড়াশোনা করে, যেমন আমরা এবার করব।
এদের সম্পর্কে আমরা যা যা শিখব সব কিছুই আসবে এই দুটো তথ্য থেকে--

- এক, সংজ্ঞাটা-- একটা real, symmetric matrix A-কে positive definite বলে, যদি

$$\forall \vec{x} \neq \vec{0} \quad \vec{x}' A \vec{x} > 0.$$

- দুই, যেকোনো positive definite matrix A-র normal form হয় identity matrix.

এবার তবে প্রথম theorem-টা দেওয়া যাক--

> **THEOREM**
>
> A real, symmetric matrix A is pd if and only if it can be written as $A = P'P$ for some nonsingular, real matrix P.

প্রমাণ অতি সহজ--যদি A-টা pd হয়, তবে A-র normal form আমরা জানি I. সুতরাং $A = P'IP$ হবে, যেখানে P' হল row operation-গুলোর matrix, আর P হল congruent column operation-গুলোর matrix. ব্যস্, তার মানে $A = P'P$. আর P যে nonsingular, সে তো জানিই। বিপরীতপক্ষে, যদি $A = P'P$ হিসেবে লেখা যায়, তবে যেকোনো $\vec{x} \neq \vec{0}$ নাও। সেক্ষেত্রে $\vec{y} = P\vec{x} \neq \vec{0}$ হবে, কারণ[5] P-টা nonsingular. অতএব

$$\vec{x}'A\vec{x} = \vec{x}'P'P\vec{x} = \vec{y}'\vec{y} = \|\vec{y}\|^2 > 0,$$

যেহেতু $\vec{y} \neq \vec{0}$. তার মানেই তো A হচ্ছে positive definite!
এই প্রমাণটা যদি বুঝে থাকো তবে নীচের অংকটা ঝট্ করে হয়ে যাবে--

Exercise 26: Prove that any real, symmetric matrix A is nonnegative definite if and only if it can be written as $A = B'B$ for some real matrix B. ■

এইবার আরেকটা সহজ theorem, যেটা অনেক কাজে লাগে--

> **THEOREM**
>
> Any positive definite matrix has determinant > 0.

প্রমাণ একলাইনে--যদি A একটা positive definite matrix হয়, তবে আমরা এক্ষুণি দেখলাম যে, $A = P'P$ হিসেবে লেখা যায়, যেখানে P হল nonsingular. সুতরাং

$$|A| = |P'P| = |P'| \cdot |P| = |P|^2 > 0,$$

যেহেতু $|P'| = |P| \neq 0$.
এবার একই রকম একটা জিনিস তুমি প্রমাণ কর দেখি nonnegative definite-দের বেলায়--

Exercise 27: Prove that if A is a nonnegative definite matrix, then $|A| \geq 0$. ■

Exercise 28: এমন একটা 2×2 সাইজের matrix দিতে পারো, যেটা pd নয়, কিন্তু যার determinant হল > 0? ■

এবার একটা জিনিস শিখব, যেটা দেখতে বেশ জটিল মনে হবে, কিন্তু সরাসরি সংজ্ঞা থেকেই প্রমাণ হয়ে যাবে, normal form-ও লাগাতে হবে না! একটা উদাহরণ দিয়ে শুরু করি।

[5] যদি $\vec{y} = \vec{0}$ হত, তবে $\vec{x} = P^{-1}\vec{y} = \vec{0}$ হত।

Example 13: ধরো একটা positive definite matrix দিলাম--

$$\begin{bmatrix} 209 & 154 & 147 & 187 \\ 154 & 198 & 156 & 199 \\ 147 & 156 & 180 & 153 \\ 187 & 199 & 153 & 247 \end{bmatrix}.$$

এবার যেকোনো কয়েকটা row নাও (যতগুলো খুশি, যা খুশি!)। ধরো আমি নিলাম দ্বিতীয় আর চতুর্থ row. এবার সেই একই column-গুলোও নাও, মানে দ্বিতীয় আর চতুর্থ column. এছাড়া বাকি সব কিছুকে মুছে ফ্যালো। তা হলে তোমার হাতে থাকবে খালি $2 \times 2 = 4$-টে entry, যাদের মোটা করে দেখিয়েছি--

$$A = \begin{bmatrix} 209 & 154 & 147 & 187 \\ 154 & \mathbf{198} & 156 & \mathbf{199} \\ 147 & 156 & 180 & 153 \\ 187 & \mathbf{199} & 153 & \mathbf{247} \end{bmatrix}$$

এদের নিয়ে একটা submatrix হয়--

$$B = \begin{bmatrix} 198 & 199 \\ 199 & 247 \end{bmatrix}.$$

দেখাও যে, এটাও একটা positive definite matrix হবে।

SOLUTION: একটা কায়দা হল এটার normal form বার করে করা। কিন্তু আমাদের বলা ছিল যে, বড় matrix-টা positive definite. এই তথ্যটা ব্যবহার করে কাজটা অনেক সহজে হয়ে যাবে। যেকোনো একটা $\vec{x} \in \mathbb{R}^2$ নাও, যেখানে $\vec{x} \neq \vec{0}$. আমরা দেখাব যে, এরকম যেকোনো $\vec{x}$-এর জন্যই $\vec{x}'B\vec{x} > 0$ হবে। এখানে $\vec{x} \in \mathbb{R}^2$ আছে, এটাকে লম্বা করে আমরা $\mathbb{R}^4$-এর একটা vector বানাব এইভাবে। যদি $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ হয় তবে আমাদের নতুন vector-টা হবে $\vec{v} = \begin{bmatrix} 0 \\ x \\ 0 \\ y \end{bmatrix}$.

লক্ষ্য কর, আমরা x আর y-কে দ্বিতীয় আর চতুর্থ স্থানে রেখেছি, এবং বাকি জায়গাগুলো 0 দিয়ে ভরে দিয়েছি। এরকম করার কারণ হল A থেকে B আমরা পেয়েছিলাম খালি দ্বিতীয় আর চতুর্থ row এবং column-গুলো নিয়ে। যেহেতু A ছিল positive definite, তাই অতি অবশ্যই $\vec{v}'A\vec{v} > 0$ হবে। এদিকে লক্ষ্য কর--

$$\vec{v}'A\vec{v} = \vec{x}'B\vec{x},$$

কারণ 0 দিয়ে যাই গুণ করবে 0 হয়ে যাবে! সুতরাং $\vec{x}'B\vec{x} > 0$, ঠিক যা চেয়েছিলাম! ■

এই যে আমরা যতগুলো খুশি এবং যা খুশি row এবং ঠিক সেই সেই column-গুলো নিয়ে একটা submatrix বানালাম, এরকম submatrix-দের বলে **principal submatrix**. সুতরাং নীচের theorem-টা নিশ্চয়ই আর প্রমাণ করে দিতে হবে না!

THEOREM
Every principal submatrix of a positive definite matrix is positive definite.

Exercise 29: বল তো, নীচের কথাগুলোর মধ্যে কোনগুলো ঠিক? ভুলগুলোর বেলায় একটা করে counterexample দিতে ভুলোনা যেন।

1. Every principal submatrix of a nonnegative definite matrix is nonnegative definite.
2. Every principal submatrix of a positive semidefinite matrix is positive semidefinite.
3. Every principal submatrix of a negative definite matrix is negative definite.
4. Every principal submatrix of an indefinite matrix is indefinite.

■

Exercise 30: ঠিক না ভুল বলো--

Every diagonal entry of a pd matrix must be positive.

■

Exercise 31: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). If the quadratic form $x^2 + \lambda(y^2 + z^2) + 2xy$ is positive definite, then

(A) $\lambda = 5$ (B) $\lambda > 1$ (C) $\lambda < 1$ (D) none of these

(2021.CC12.1j)
HINT: প্রথমে diagonal entry-গুলো নিয়ে চিন্তা করো। তারপর 2×2 সাইজের principal submatrix-গুলোর determinant নিয়ে। ■

58.5 Hessian matrix (আবার!)

আমরা এই অধ্যায়ে আগেই বলেছিলাম যে, calculus-এ second derivative দিয়ে যা যা কাজ হয়, মোটামুটিভাবে সেই একই কাজ multivariable calculus-এর বেলায় হয় Hessian দিয়ে। এবার সেটা নিয়েই সংক্ষেপে আলোচনা করব। এ বিষয়ে আরো তথ্য পাবে আমাদের "বাংলায় বোঝানো ইংরাজী বই" সিরিজের Multivariable Calculus বইতে।
ধরো একটা function আছে $f : \mathbb{R}^2 \to \mathbb{R}$. এর গ্রাফটা একটা surface-এর মত দেখতে (Fig 28)। কয়েকটা জায়গায় পাহাড়ের মত দেখতে পাচ্ছ? ওদের চূড়াগুলোকে বলে একেকটা **local maximum**. আবার কয়েক জায়গায় খাদের মত আছে, ওদের গভীরতম বিন্দুগুলো হল একেকটা **local minimum**. এদেরকে ছবি থেকে বোঝা সহজ। অনেক সময়ে এ কাজটা আমরা অংক কষে করতে চাই। তখন আমরা multivariable calculus ব্যবহার করি। তার জন্য প্রথমে **tangent plane**-এর ধারণাটা বোঝা দরকার। আমরা জানি যে, পৃথিবীটা গোল। অথচ পায়ের তলার জমি দিব্যি সমতল বলেই বোধ হয়। তার কারণ পৃথিবীর খুব অল্প অংশই আমাদের চোখের পাল্লায় ধরা পড়ে। তাই ভূপৃষ্ঠের বক্রতাটা মালুম হয় না। ঠিক

Fig 28

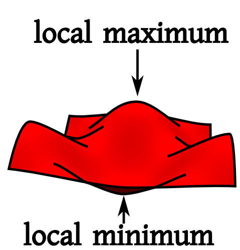
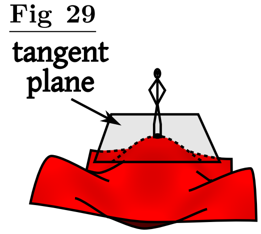


Fig 29



একই ব্যাপার হয় আমাদের surface-টার বেলাতেও। যদি ওর উপর কোনো বিন্দুতে একটা লোক দাঁড়িয়ে থাকে (Fig 29), যে খুব বেশী দূর অবধি দেখতে পায় না, তবে তার কাছে মনে হবে যেন সে একটা plane-এর উপর দাঁড়িয়ে আছে। সেই plane-টাকেই বলে ওই বিন্দুতে surface-টার tangent plane. এই tangent plane-টা যে একটা horizontal plane হবেই, এমন কোনো কথা নেই। কিন্তু যদি লোকটা কোনো local maximum বা local minimum বিন্দুতে দাঁড়িয়ে থাকে (যেমন Fig 29-এর বেলায় আছে), তবে tangent plane-টা horizontal হতে বাধ্য। তাই local maximum বা local minimum বার করার প্রথম ধাপ হল সেই সব বিন্দু খুঁজে বার করা, যেখানে tangent plane-টা horizontal হয়। সেটা অংক কষে বোঝার কায়দা হল $\nabla f \equiv \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ বার করা। যেখানে tangent plane-টা সমতল হয়, সেখানে $\nabla f = \vec{0}$ হবে। এইসব বিন্দুকে বলে একেকটা **stationary point**. দুঃখের কথা যে, local maximum আর local minimum-গুলো একেকটা stationary point হলেও, এর উল্টো দিকটা ঠিক নয়, মানে এমন stationary point সম্ভব, যেখানে local maximum বা minimum কোনোটাই নেই। সেই সব বিন্দুকে বলে **saddle point**. সুতরাং যেসব বিন্দুতে $\nabla f = \vec{0}$ হয়েছে, সেগুলো বার করার পর আমাদের কাজ হল তাদেরকে তিনটে ভাগে ভাগ করা--

- এক, যেখানে local maximum আছে,
- দুই, যেখানে local minimum আছে,
- আর তিন, যেখানে কোনোটাই নেই, মানে saddle point আছে।

এই কাজেই Hessian matrix ব্যবহার করা হয়--

$$H(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}.$$

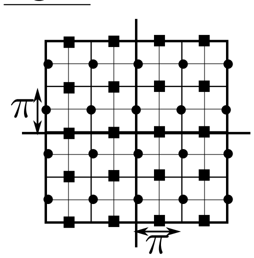
এখানে আমরা ধরে নেব যে $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. ফলে Hessian-টা একটা symmetric matrix হবে। লক্ষ্য করো, এটা (x,y)-এর function, মানে বিভিন্ন জায়গায় Hessian-টা বিভিন্ন হতে পারে। তোমার কাজ হবে ওই যে বিন্দুগুলো পেয়েছিলে, তার প্রত্যেকটাতে Hessian বার করা। যেখানে যেখানে Hessian-টা pd হবে, সেখানে local minimum থাকবে, যেখানে nd হবে, সেখানে local maximum, আর যেখানে id হবে সেখানে saddle point. স্বাভাবিকভাবেই প্রশ্ন উঠবে যে, Hessian-টা এই তিন ধরণের বাইরে অন্য কিছু হলে (যেমন psd বা nsd হলে) কী বলা যায়। উত্তর হল, সেক্ষেত্রে কিছুই বলা যায় না। কয়েকটা উদাহরণ দেখে পুরো জিনিসটা হজম করা যাক।

Example 14: ধরো $f(x,y) = \cos(x)\sin(y)$. এর local maximum আর minimum-গুলো বার করো। সেই সঙ্গে কোনো saddle point থাকলে সেগুলোও বলো।

SOLUTION: এখানে $\nabla f = (-\sin x \sin y, \cos x \cos y)$.
এটা কোথায় কোথায় $\vec{0}$ হবে, সেটা Fig 30-এ দেখিয়েছি--

- এক, যখন $\sin x = 0$ আর $\cos y = 0$ হবে (ছবিতে এদের কালো গোল দিয়ে এঁকেছি),

Fig 30



- আর দুই, যখন $\cos x = 0$ আর $\sin y = 0$ হবে (ছবিতে এরা কালো square)।

এদের শ্রেণীবিভাগ করার জন্য Hessian-এর ফর্মুলাটা বার করে রাখি--

$$H(x,y) = \begin{bmatrix} -\cos x \sin y & -\sin x \cos y \\ -\sin x \cos y & -\cos x \sin y \end{bmatrix}.$$

এবার বিভিন্ন stationary point-এ Hessian বার করব--

- যখন $\sin x = 0$ আর $\cos y = 0$ হবে, তখন Hessian-টা দাঁড়াবে $\begin{bmatrix} -\cos x \sin y & 0 \\ 0 & -\cos x \sin y \end{bmatrix}$. এদিকে $\cos x = \pm 1$ এবং $\sin y = \pm 1$. তাই $\cos x \sin y = \pm 1$. তার মানে দুটো কেস হচ্ছে--
  - যদি $\cos x \sin y = 1$ হয় তবে Hessian-টা হচ্ছে $-I$, যেটা nd. তার মানে local maximum হচ্ছে যখন $x = 2m\pi$ আর $y = \frac{\pi}{2} + 2n\pi$ চেহারার (যেখানে $m, n \in \mathbb{Z}$.)।
  - যদি $\cos x \sin y = -1$ হয় তবে Hessian-টা হচ্ছে I, যেটা pd. তার মানে local minimum হচ্ছে যখন $x = \pi + 2m\pi$ আর $y = \frac{\pi}{2} + 2n\pi$ চেহারার (যেখানে $m, n \in \mathbb{Z}$.)।
- যখন $\cos x = 0$ আর $\sin y = 0$ হবে, তখন Hessian-টা হবে $\begin{bmatrix} 0 & -\sin x \cos y \\ -\sin x \cos y & 0 \end{bmatrix}$. এখানেও দুটো কেস--
  - যখন $\sin x \cos y = 1$, তখন Hessian-টা হবে $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$, যেটা id.
  - যখন $\sin x \cos y = -1$, তখন Hessian-টা হবে $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, যেটাও id.

  তার মানে saddle point হচ্ছে যখন $x = \frac{\pi}{2} + m\pi$ আর $y = n\pi$ চেহারার (যেখানে $m, n \in \mathbb{Z}$.)।

এই শ্রেণীবিভাগটা দেখিয়েছি Fig 31-এ। এখানে ভরাট কালো গোল মানে হল local maximum, ফাঁকা গোল মানে local minimum, আর ভরাট কালো square মানে saddle point. ■

Exercise 32: Let $f(x,y) = x^2 + y^2 + xy$. Find the Hessian matrix of f at $(0,0)$ and show that f has a local minimum at the origin.[3] **(2020.CC12.3bii)**
HINT: ছবিটা দেখিয়েছি Fig 32-এ। ■

Fig 31

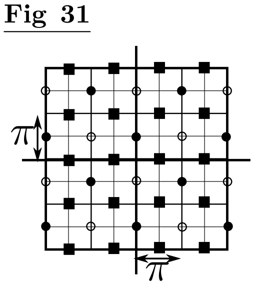


Fig 32



আগেই বলেছি যে, Hessian-টা যদি pd, nd বা id কোনোটাই না হয়, তবে কিছুই বলা যায় না। সেরকম কিছু উদাহরণ দেখার আগে বলে নিই, pd, nd, id ক্ষেত্রে এই কায়দাটা কেন কাজ করে। এটা আসে Taylor approximation থেকে। ধারণাটা আমরা এই অধ্যায়ে আগেই উল্লেখ করেছি। জিনিসটাকে গুছিয়ে এভাবে লেখা যায়--

$$f(\vec{x}) \approx f(\vec{a}) + \nabla f(\vec{a})(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})'H(\vec{a})(\vec{x}-\vec{a}).$$

Stationary point-গুলোতে $\nabla f(\vec{a}) = \vec{0}$ হয়, তাই তখন--

$$f(\vec{x}) \approx f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})'H(\vec{a})(\vec{x}-\vec{a}).$$

লক্ষ্য করো এটা প্রায় একটা quadratic form. "প্রায়" বললাম, কারণ শুরুতে একটা constant যোগ হয়ে আছে। কিন্তু তাতে তো আর surface-এর আকারটা বদলাবে না, খালি বড়জোর একটু উঠবে বা নামবে, এই যা। তাই quadratic form-টার যা আচরণ (local maximum, minimum বা saddle point), সেই একই আচরণ তুমি ওই (a, b) বিন্দুর ধারেকাছে দেখতে পাবে $f(x, y)$-এর বেলাতেও। সেই জন্যই Hessian-এর কায়দাটা কাজ করে।
এইবার বলি Hessian-টা যখন psd বা nsd হয়, তখন কী হয়। তার জন্য নীচের অংকগুলো করো। এখানে প্রতিক্ষেত্রেই Hessian-টা হবে $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$, যেটা একটা psd matrix. কিন্তু $f(x, y)$-এর আচরণ এেকেকক্ষেত্রে এেকেকরকম।

Exercise 33: In each case below $(0, 0)$ is a stationary point, and the Hessian there is $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$. Determine if we have a local maximum, minimum or a saddle point at $(0, 0)$ in each case.

1. $f(x, y) = x^2 + y^4$.
2. $f(x, y) = x^2 - y^4$.

■

DAY 59
Bilinear form (part 1)

এবার আমরা bilinear form বলে একটা ধারণা নিয়ে আলোচনা করব।

59.1 কী জিনিস?

একটা উদাহরণ নিয়ে শুরু করি।

Example 15: মনে করো $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ দিলাম $f(x, y) = xy$. এটা কি একটা linear transformation? যদি $y = 5$ বসাই, তবে খালি x-এর function হিসেবে কি এটা একটা linear transformation? আবার যদি $x = -1$ বসাতাম, তবে খালি y-এর function হিসেবে কি এটা একটা linear transformation হত?
SOLUTION: যদি x, y দুজনেই একসঙ্গে বদলায়, তবে $f(x, y)$ মোটেই linear transformation নয়, কারণ $f(2(1, 1)) = f(2, 2) = 4 \neq 2f(1, 1)$.
কিন্তু যদি $y = 5$ বসাই, তবে খালি x-এর function হিসেবে পাব $f(x, 5) = 5x$, যেটা অবশ্যই একটা linear transformation. একইভাবে $x = -1$ বসালে খালি y-এর function হিসেবে পাব $f(-1, y) = -y$, যেটাও একটা linear transformation. ■

এইটা হল একটা bilinear form-এর উদাহরণ। এরকম function-রা দুটো vector-কে মিলিয়ে একটা scalar বানায়। যদি কোনো একটাকে স্থির রেখে খালি অন্যটার function হিসেবে দেখি, তবে linear transformation হবে। ঠিক যেন একটা function-এর মধ্যে দুটো linear transformation পাশাপাশি অবস্থান করছে। সেই কারণেই bi-linear শব্দটা এসেছে।

> DEFINITION: **Bilinear form**
>
> Let V be a vector space over a field F. By a **bilinear form** on V we mean a function $f : V \times V \to F$ such that for each fixed $\vec{b} \in V$ the function $\vec{x} \mapsto f(\vec{x}, \vec{b})$ is linear, and for each $\vec{a} \in V$ the function $\vec{y} \mapsto f(\vec{a}, \vec{y})$ is also linear.

কয়েকটা উদাহরণ দেখে সংজ্ঞাটা সড়গড় হয়ে নেওয়া যাক।

Example 16: নীচে কয়েকটা $f : V \times V \to F$ দিলাম, এদের মধ্যে কারা bilinear form বলো।

1. $F = \mathbb{R},\ V = \mathbb{R}^2.\ f((a, b), (c, d)) = a - b + c - d$
2. $F = \mathbb{R},\ V = \mathbb{R}^2.\ g((a, b), (c, d)) = ad + bc$

SOLUTION: ধরো প্রথম ক্ষেত্রে (c, d) স্থির রাখলাম। তবে কেবলমাত্র (a, b)-র function হিসেবে ভাবলে $f((a, b), (c, d))$ হবে $a - b +$constant, কারণ $c - d$ হল একটা constant. প্রশ্ন হল এটা কি (a, b)-র একটা linear transformation? আমরা জানি (a, b)-র linear transformation-রা দেখতে হয় এইরকম--

কোনো সংখ্যা$\times a$ + কোনো সংখ্যা$\times b$.

আমাদের বেলায় আছে $a - b +$constant. ওই শেষের constant-টাই সমস্যা করে দিয়েছে। অবশ্য যদি constant-টা শূন্য হয়, তবে কোনো সমস্যা নেই। কিন্তু এখানে constant-টা হল $c - d$, যেটার শূন্য হবার কোনো গ্যারান্টি নেই, যেমন $c = 2$, $d = 1$ নিলেই constant-টা হবে $c - d = 1$. সুতরাং $f((a, b), (2, 1))$ মোটেই (a, b)-র linear transformation হচ্ছে না। সুতরাং $f((a, b), (c, d))$-টা bilinear form নয়।
দ্বিতীয় ক্ষেত্রে যদি (a, b) স্থির রাখি, তবে কেবলমাত্র (c, d)-র function হিসেবে--

$g((a, b), (c, d)) =$ কোনো সংখ্যা$\times d$ + কোনো সংখ্যা$\times c$,

যেটা অবশ্যই (c, d)-র একটা linear transformation. আবার যদি (c, d) স্থির রাখতাম, তবে (a, b)-র function হিসেবে--

$g((a, b), (c, d)) =$ কোনো সংখ্যা$\times a$ + কোনো সংখ্যা$\times b$,

যেটা (a, b)-র একটা linear transformation. সুতরাং সব মিলিয়ে বলতে পারি যে, $g((a, b), (c, d))$ হল একটা bilinear form. ■

Exercise 34: নীচে কয়েকটা $f : V \times V \to F$ দিলাম, এদের মধ্যে কারা bilinear form বলো।

1. $F = \mathbb{R},\ V = \mathbb{R}^2.\ f((a, b), (c, d)) = a + b + c + d$
2. $F = \mathbb{R},\ V = \mathbb{R}^2.\ f((a, b), (c, d)) = ad - bc$
3. $F = \mathbb{R},\ V = \mathbb{R}^2.\ f((a, b), (c, d)) = ab - cd$

4. $F = \mathbb{R},\ V = \mathbb{R}^2.\ f((a,b),(c,d)) = 2ac + 3ad - bc + 5bd$
5. $F = \mathbb{R},\ V = \mathbb{C}.\ f(w,z) = w\overline{z}$
6. $F = \mathbb{R},\ V = \mathbb{C}.\ f(w,z) = wz$

■

Exercise 35: যদি $V = \mathbb{R}^2$ এবং $F = \mathbb{R}$ হয়, তবে দেখাও যে, যেকোনো $\alpha, \beta, \gamma, \delta \in \mathbb{R}$-এর জন্যই $f((x,y),(w,z)) = \alpha xw + \beta xz + \gamma yw + \delta yz$ হল একটা bilinear form. ■

Exercise 36: ধরো $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ হল একটা bilinear form. যদি $\alpha = f\big((1,0),(1,0)\big)$, $\beta = f\big((1,0),(0,1)\big)$, $\gamma = f\big((0,1),(1,0)\big)$ এবং $\delta = f\big((0,1),(0,1)\big)$ হয়, তবে দেখাও যে, যেকোনো $(x_1,x_2),(y_1,y_2)$-র জন্যই

$$f\big((x_1,x_2),(y_1,y_2)\big) = \alpha x_1y_1 + \beta x_1y_2 + \gamma x_2y_1 + \delta x_2y_2$$

হবে। ■

এই দুটো অংক করার পর নীচের অংকটা সহজ হওয়া উচিত।

Exercise 37: এমন কোনো bilinear form দিতে পারো, $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, যাতে $f\big((x_1,x_2),(y_1,y_2)\big)$-টা x_iy_j-দের একটা linear combination না হয়? ■

Exercise 38: এমন কোন্ কোন্ bilinear form আছে, যারা linear transformation-ও বটে? ■

59.2 Matrix representation

Finite dimensional vector space-দের ক্ষেত্রে কোনো linear transformation বা quadratic form-কে যেমন কোনো basis-এর সাপেক্ষে একটা matrix দিয়ে প্রকাশ করা যায়, bilinear form-দের বেলাতেও তাই। Linear transformation-দের বেলায় চেহারাটা ছিল $f(\vec{x}) = A\vec{x}$, (যেখানে A যেকোনো matrix), আর quadratic form-এর বেলায় $q(\vec{x}) = \vec{x}'A\vec{x}$ (যেখানে A যেকোনো real, symmetric matrix)। একইভাবে bilinear form-দের বেলায় চেহারাটা হবে $f(\vec{x},\vec{y}) = \vec{x}'A\vec{y}$, যেখানে A যেকোনো square matrix. এরকম চেহারার যেকোনো function-ই যে একটা bilinear form হবে, সেটা নীচের অংকটা দেখলেই বুঝবে।

Example 17: যদি A হয় $n \times n$ সাইজের যেকোনো matrix, তবে দেখাও যে $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ একটা bilinear form, যেখানে $f(\vec{x},\vec{y}) = \vec{x}'A\vec{y}$.

SOLUTION: যদি $\vec{y}$-কে স্থির রাখি, ধরো $\vec{y} = \vec{b}$, তবে খালি $\vec{x}$-এর function হিসেবে পাব $g(\vec{x}) = f(\vec{x},\vec{b}) = \vec{x}'A\vec{b}$. দেখাতে হবে যে, এটা একটা linear transformation. তার জন্য যেকোনো $\alpha, \beta \in \mathbb{R}$ এবং $\vec{x}_1, \vec{x}_2 \int \mathbb{R}^n$ নাও। তবে

$$g(\alpha\vec{x}_1 + \beta\vec{x}_2) = (\alpha\vec{x}_1 + \beta\vec{x}_2)'A\vec{b} = \alpha\vec{x}_1'A\vec{b} + \beta\vec{x}_2'A\vec{b} = \alpha g(\vec{x}_1) + \beta g(\vec{x}_2).$$

তাই $g(\vec{x})$ একটা linear transformation.
একই রকম যুক্তিতে যদি $\vec{x} = \vec{a}$ স্থির রাখি, তবে $h(\vec{y}) = f(\vec{a},\vec{y}) = \vec{a}'A\vec{y}$-ও একটা linear transformation হবে।
তাই $f(\vec{x},\vec{y})$ সত্যিই একটা bilinear form হচ্ছে। ■

এর উল্টো দিকের প্রমাণটা সামান্য বেশী জটিল--কোনো finite dimensional vector space-এর ক্ষেত্রে যেকোনো bilinear form-এরই যে এই চেহারা হতে বাধ্য। সেটা নিয়েই নীচের theorem-টা--

Existence

Let V be a finite dimensional vector space over a field F. Let $f : V \times V \to F$ be any bilinear form. Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be any basis of V. Then show that there must exist a matrix $A_{n\times n}$ such that if $\vec{x}, \vec{y} \in V$ have coordinates $(x_1, ..., x_n)$ and $(y_1, ..., y_n)$, respectively, wrt this basis, then

$$f(\vec{x}, \vec{y}) = (x_1, ..., x_n) A \left[\begin{array}{c} y_1 \\ \vdots \\ y_n \end{array}\right].$$

Proof: প্রমাণটা notation-এর দিক দিয়ে একটু জটিল লাগতে পারে। মূল ধারণাটা প্রথমে একটা উদাহরণ দেখে নিই। ধরো $f\big((x_1, x_2), (y_1, y_2)\big) = x_1y_2 - x_2y_1$. আর আমাদের basis-টা হল সহজতম basis, মানে canonical basis. আমরা এমন একটা matrix খুঁজছি, $A = \left[\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array}\right]$, যাতে

$$\forall (x_1, x_2), (y_1, y_2) \in \mathbb{R}^2 \quad x_1y_2 - x_2y_1 = \left[\begin{array}{cc} x_1 & x_2 \end{array}\right] A \left[\begin{array}{c} y_1 \\ y_2 \end{array}\right]$$

হয়। যদি এরকম একটা matrix সত্যিই থাকতে হয়, তবে তাকে খুঁজে পাওয়ার জন্য $(x_1, x_2) = (1, 0)$ আর $(y_1, y_2) = (1, 0)$ বসালেই পাবে--

$$1 \times 0 - 0 \times 1 = \left[\begin{array}{cc} 1 & 0 \end{array}\right] \left[\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array}\right] \left[\begin{array}{c} 1 \\ 0 \end{array}\right].$$

এ থেকে a_{11} বেরিয়ে যাবে--

$$0 = a_{11}.$$

একইভাবে $(x_1, x_2) = (1, 0)$ রেখে $(y_1, y_2) = (0, 1)$ করলে বেরোবে $a_{12} = 1$. এইভাবে অন্য entry-গুলোও বেরিয়ে গেলে পাবে $A = \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right]$. এতে বোঝা গেল যে, যদি এমন কোনো A সত্যিই থাকে, যেটা (*) মেনে চলে, তবে সেটা এই A-টা না হয়ে যায় না। সত্যিই যে এই A দিয়ে কাজ হবে, সেটা এবার যেকোনো (x_1, x_2), (y_1, y_2) নিয়ে (*)-এ বসিয়ে দেখলেই প্রমাণ হয়ে যাবে।

এই একই পথে আমরা theorem-এর প্রমাণটা করব যেকোনো basis-এর ক্ষেত্রেই। আমাদের basis-টা দেওয়া ছিল $\{\vec{v}_1, ..., \vec{v}_n\}$. তাই matrix-টার (i, j)-th entry বার করার জন্য $\vec{x} = \vec{v}_i$ আর $\vec{y} = \vec{v}_j$ বসাব। তাহলেই পাবে $f(\vec{v}_i, \vec{v}_j)$. সুতরাং matrix-টার (i, j)-th entry যে $f(\vec{v}_i, \vec{v}_j)$ হবে, সেটা বুঝে গেলাম। এবার খালি দেখানোর অপেক্ষা যে, এই matrix-টা যেকোনো $\vec{x}$, $\vec{y}$-এর জন্যই কাজ করবে।

Let $A_{n\times n} = ((a_{ij}))$, where $a_{ij} = f(\vec{v}_i, \vec{v}_j)$.

Then

$$\begin{aligned} f(\vec{x},\vec{y}) &= f\left(\sum_i x_i\vec{v}_i, \sum_j y_j\vec{v}_j\right) \\ &= \sum_i x_i f\left(\vec{v}_i, \sum_j y_j\vec{v}_j\right) \quad [\because f \text{ is a bilinear form}] \\ &= \sum_i\sum_j x_i y_j f(\vec{v}_i,\vec{v}_j) \quad [\because f \text{ is a bilinear form}] \\ &= \sum_i\sum_j x_i y_j a_{ij} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \end{aligned}$$

as required.

[Q.E.D]

আমরা জানলাম যে, কোনো একটা basis দেওয়া থাকলে যেকোনো bilinear form-কে তার সাপেক্ষে একটা matrix দিয়ে প্রকাশ করা যায়। এই matrix-টা আবার unique-ও হয়। মানে একই basis-এর সাপেক্ষে একই bilinear form-এর জন্য যদি একজন বলে matrix-টা A এবং আরেকজন বলে matrix-টা B, তবে $A = B$ হতে বাধ্য। এই কথাটাকে একটা theorem-এর আকারে লিখে প্রমাণ করে রাখা যাক। বারবার কাজে লাগবে।

Uniqueness

Let $A_{n\times n}$ and $B_{n\times n}$ be two matrices such that

$$\forall \vec{x},\vec{y} \in \mathbb{R}^n \quad \vec{x}'A\vec{y} = \vec{x}'B\vec{y}.$$

Then $A = B$.

Proof: এটা প্রমাণ করার জন্য আমাদের সুপরিচিত canonical basis-টার কথা ভাবো--

Let $\{\vec{e}_1, ..., \vec{e}_n\}$ be the canonical basis of $\mathbb{R}^n$.
Then, by the given condition, we have

$$\forall i,j \quad \vec{e}_i'A\vec{e}_j = \vec{e}_i'B\vec{e}_j. \qquad (*)$$

এটা পেলাম $\vec{x}$-এর জায়গায় $\vec{e}_i$, এবং $\vec{y}$-এর জায়গায় $\vec{e}_j$ বসিয়ে। এবার লক্ষ্য করো--

But $\vec{e}_i'A\vec{e}_j$ is the (i,j)-th entry of A, and $\vec{e}_i'B\vec{e}_j$ is the (i,j)-th entry of B.

এটা যদি স্পষ্ট না হয়, তবে একটা উদাহরণ নিয়ে দেখো। মনে করো $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. তবে গুণটা করলেই দেখবে--

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2.$$

Hence (*) means $A = B$.

[Q.E.D]

Exercise 39: নীচের প্রতিক্ষেত্রে একটা করে bilinear form এবং একটা করে basis দেওয়া আছে। ওই basis-এর সাপেক্ষে bilinear form-টার matrix বার করো।

1. $f\big((x_1,x_2),(y_1,y_2)\big) = x_1y_2 + x_2y_1$, যেখানে basis-টা হল canonical basis.
2. $f\big((x_1,x_2),(y_1,y_2)\big) = x_1y_2 + x_2y_1$, যেখানে basis-টা হল $\{(1,0),(1,1)\}$.

■

Exercise 40: যদি $f(\vec{x},\vec{y})$ একটা bilinear form হয়, তবে $q(\vec{x}) = f(\vec{x},\vec{x})$ হবে একটা quadratic form. যদি কোনো basis-এর সাপেক্ষে f-এর matrix-টা A হয়, তবে সেই একই basis-এর সাপেক্ষে q-এর matrix-টা কী হবে? ■

নীচের প্রশ্ন কয়টা তোমাকে বোকা বানানোর জন্য। প্যাঁচটা ধরে ফেলতে পারলে মুহূর্তের মধ্যে হয়ে যাবে।

Exercise 41: "যদি A একটা PD matrix হয়, তবে $f(\vec{x},\vec{y}) = \vec{x}'A\vec{y}$ কখনোই negative হতে পারে না"--ঠিক না ভুল? ■

Exercise 42: $\mathbb{R}^2$-র উপরে একটা bilinear form আছে, $f(\vec{x},\vec{y})$, যেটা কখনোই negative হয় না। বলতে পারো $f((1,1),(3,5)) = ?$ ■

Exercise 43: $\mathbb{R}^3$-র উপরে এমন একটা bilinear form আছে, $f(\vec{x},\vec{y})$, যার জন্য

$$\forall \vec{x},\vec{y} \in \mathbb{R}^3 \quad |f(\vec{x},\vec{y})| \leq 10.$$

বলতে পারো $f((1,1,1),(3,5,0)) = ?$ ■

Exercise 44: ঠিক না ভুল বলো--

1. $\mathbb{R}$-এর উপরে যেকোনো inner product-ই একটা bilinear form.
2. $\mathbb{C}$-এর উপরে যেকোনো inner product-ই একটা bilinear form.

■

DAY 60
Bilinear form (part 2)

60.1 আবার equivalence

গতকাল আমরা দেখেছি যে, basis বদলালে একই bilinear form-এর বিভিন্ন matrix পাওয়া যায়। এখানে ফের একটা equivalence-এর গন্ধ পাচ্ছ নিশ্চয়ই? দুটো matrix-কে আমরা equivalent বলব, যদি ওরা দুজনেই একই bilinear form-কে প্রকাশ করে (দুটো basis-এর সাপেক্ষে)।

Exercise 45: এটা যে সত্যিই একটা equivalence, সেটা পরীক্ষা করে দেখে নিও, মানে সেই তিনটে শর্ত, reflexive, symmetric এবং transitive. তিনটেই অবশ্য চোখে দেখেই বোঝা যায় (সত্যিই যায় তো?)। ■

Linear transformation-এর বেলায় equivalence-টা ছিল similarity. আবার quadratic form-এর বেলায় সেটা ছিল congruence. এখানে কি আবার নতুন কিছু একটা হবে? না, নীচের অংকটা আমাদের অভয় দিয়ে বলছে যে, congruence দিয়েই এখানেও কাজ চলে যাবে।

Example 18: Prove that any two matrix representations of a bilinear form are congruent.[2]

(2020.CC12.3bi)

SOLUTION:

> Let $f(\vec{x},\vec{y})$ be a bilinear form on a vector space, V. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases of V.
>
> Let the matrix representations of f wrt $\mathcal{B}_1, \mathcal{B}_2$ be A_1, A_2, respectively.

এর মানে কী, সেটা বিশদভাবে লিখে নেওয়া যাক--

> Thus, if we take any $\vec{x},\vec{y}\in V$ with coordinates $(p_1,...,p_n)$ and $(q_1,...,q_n)$, respectively, wrt $\mathcal{B}_1$, and $(\tilde{p}_1,...,\tilde{p}_n)$ and $(\tilde{q}_1,...,\tilde{q}_n)$, respectively, wrt $\mathcal{B}_2$, then
>
> $$f(\vec{x},\vec{y}) = \begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix} A_1 \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} = \begin{bmatrix} \tilde{p}_1 & \cdots & \tilde{p}_n \end{bmatrix} A_2 \begin{bmatrix} \tilde{q}_1 \\ \vdots \\ \tilde{q}_n \end{bmatrix}.$$

আমরা জানি যে $(p_1,...,p_n)$ এবং $(\tilde{p}_1,...,\tilde{p}_n)$-র মধ্যে একটা সম্পর্ক আছে--

> $$\begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} = C \begin{bmatrix} \tilde{p}_1 \\ \vdots \\ \tilde{p}_n \end{bmatrix},$$
>
> where C is the change of basis matrix between $\mathcal{B}_1$ and $\mathcal{B}_2$.

C ব্যবহার হবে অন্য ক্ষেত্রটাতেও--

> Similarly,
>
> $$\begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} = C \begin{bmatrix} \tilde{q}_1 \\ \vdots \\ \tilde{q}_n \end{bmatrix}.$$

So

$$\begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix} A_1 \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} = \begin{bmatrix} \tilde{p}_1 & \cdots & \tilde{p}_n \end{bmatrix} C' A_1 C \begin{bmatrix} \tilde{q}_1 \\ \vdots \\ \tilde{q}_n \end{bmatrix}.$$

তার মানে $\mathcal{B}_2$-র সাপেক্ষে আমাদের bilinear form-টার দুটো matrix পাওয়া যাচ্ছে, A_2 আর $C'A_1C$.

$$\begin{bmatrix} \tilde{p}_1 & \cdots & \tilde{p}_n \end{bmatrix} C' A_1 C \begin{bmatrix} \tilde{q}_1 \\ \vdots \\ \tilde{q}_n \end{bmatrix} = \begin{bmatrix} \tilde{p}_1 & \cdots & \tilde{p}_n \end{bmatrix} A_2 \begin{bmatrix} \tilde{q}_1 \\ \vdots \\ \tilde{q}_n \end{bmatrix}.$$

যেহেতু এটা যেকোনো $(\tilde{p}_1, ..., \tilde{p}_n)$ আর $(\tilde{q}_1, ..., \tilde{q}_n)$-এর জন্য খাটে, তাই ভিতরের matrix দুটো সমান হতে বাধ্য--

Since this holds for all $(\tilde{p}_1, ..., \tilde{p}_n)$ and $(\tilde{q}_1, ..., \tilde{q}_n)$, hence $A_2 = C'A_1C$.
Thus, A_1 and A_2 are congruent each other, as C is nonsingular.

■

60.2 দুটো বিশেষ ধরণের bilinear form

DEFINITION: **Symmetric/alternating bilinear form**

- A bilinear form f is called **symmetric** if $\forall \vec{x}, \vec{y} \quad f(\vec{x}, \vec{y}) = f(\vec{y}, \vec{x})$.
- A bilinear form f is called **alternating** if $\forall \vec{x}, \vec{y} \quad f(\vec{x}, \vec{y}) = -f(\vec{y}, \vec{x})$.

Example 19: নীচের bilinear form-গুলোর মধ্যে কে বা কারা symmetric, alternating বা কোনোটাই নয়, বলো।

1. $f\big((x_1, x_2), (y_1, y_2)\big) = x_1 y_1$
2. $g\big((x_1, x_2), (y_1, y_2)\big) = x_1 y_2$
3. $h\big((x_1, x_2), (y_1, y_2)\big) = x_1 y_1 - x_2 y_2$
4. $k\big((x_1, x_2), (y_1, y_2)\big) = x_1 y_2 - x_2 y_1$

SOLUTION:(x_1, x_2) আর (y_1, y_2)-র স্থানবিনিময় করে দেখলেই বুঝবে যে, f আর h হল symmetric, এবং k হল alternating. কিন্তু g কোনোটাই নয়। ■

Exercise 46: নীচের bilinear form-গুলোর মধ্যে কে বা কারা symmetric, alternating বা কোনোটাই নয়, বলো।

1. $f\big((x_1,x_2),(y_1,y_2)\big)=x_1y_2+x_2y_1$
2. $g\big((x_1,x_2),(y_1,y_2)\big)=x_1y_1+x_2y_2$
3. $h\big((x_1,x_2),(y_1,y_2)\big)=x_2y_2-x_1y_1$
4. $k\big((x_1,x_2),(y_1,y_2)\big)=x_1y_1+x_1y_2+x_2y_1$

■

Example 20: Prove that a symmetric bilinear form has a symmetric matrix wrt every basis.

SOLUTION:

> Let us fix some basis. Let the matrix of a symmetric bilinear form wrt this basis be A. Shall show that $A=A'$.
> By symmetry, we have
> $$\forall \vec{x},\vec{y}\in F^n\quad \vec{x}'A\vec{y}=\vec{y}'A\vec{x}.$$
> But we also have $\vec{x}'A\vec{y}=(\vec{x}'A\vec{y})'=\vec{y}'A'\vec{x}$.

এর কারণ $\vec{x}'A\vec{y}$ হল একটা 1×1 সাইজের matrix. তাই এটা এর নিজের transpose-এর সমান।

> Hence $\forall \vec{x},\vec{y}\in F^n\quad \vec{y}'A'\vec{x}=\vec{y}'A\vec{x}$.

এখানে যে আমরা uniqueness theorem-এর যুক্তিটা লাগালাম, সেটা আশা করি বুঝতে পেরেছ।

> So $A=A'$, as required.

■

নীচের অংকটা একইভাবে করতে পারা উচিত।

Exercise 47: Prove that an alternating bilinear form has a skew-symmetric matrix wrt every basis.

■

60.3 সব bilinear form-দের vector space

যদি তুমি দুটো bilinear form-কে যোগ করো, তবে ফের একটা bilinear form-ই পাবে। সেটা বোঝা শক্ত নয়। যদি কোনো bilinear form-কে একটা scalar দিয়ে গুণ করো, তবে গুণফলটাও bilinear form হবে। এই দুই মিলিয়ে নিশ্চয়ই নীচের theorem-টা আন্দাজ করতে পারছ--

> THEOREM
> The space of all bilinear forms on a vector space is a vector space.

এবার এই vector space-টাকে নিয়ে আলোচনা করব।

60.3.1 Dimension এবং basis

বলতে পারো এই vector space-টার dimension কত হবে? অবশ্যই dimension-এর আলোচনার আগে ধরে নিতে হবে যে, যে vector space-এর উপর bilinear form নিয়ে কাজ করছি, সেটা নিজে finite dimensional. আমরা আগেই দেখেছি যে, একটা n-dimensional vector space-এর উপরে কোনো bilinear form-কে একটা $n \times n$ সাইজের matrix দিয়ে প্রকাশ করা যায়, এবং বিপরীতপক্ষে এরকম যেকোনো matrix-ই একটা bilinear form-এর জন্ম দেয়। আবার যদি দুটো bilinear form-কে যোগ করো, তবে ওদের matrix দুটোও যোগ হয়ে যায়। যদি একটা bilinear form-কে কোনো scalar দিয়ে গুণ করো, তবে তার matrix-টাও সেই scalar দিয়ে গুণ হয়ে যায়। তাই যাবতীয় $n \times n$ সাইজের matrix-এর vector space আর যাবতীয় bilinear form-এর vector space হল isomorphic. তাই ওদের dimension-ও সমান। এখন $n \times n$ সাইজের matrix-এ ঠিক n^2 খানা entry থাকে, যারা প্রত্যেকেই free[6] তাই dimension হবে n^2.

এই গেলে dimension-এর কথা। এবার basis বার করার চেষ্টা করা যাক। যদি বলি $\mathbb{R}^2$-র উপরে যত bilinear form আছে, তাদের vector space-এর একটা basis দিতে, তবে তোমাকে $2^2 = 4$-টা bilinear form দিতে হবে। এখানেও matrix দিয়ে ভাবলে সুবিধা হবে। যাবতীয় 2×2 সাইজের matrix-এর vector space-এর basis হল

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

এই matrix-গুলোর জন্য যে bilinear form-গুলো পাবে, তারা হল--

$$\begin{aligned}
f_1\big((x_1,x_2),(y_1,y_2)\big) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_1 \\
f_2\big((x_1,x_2),(y_1,y_2)\big) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_2 \\
f_3\big((x_1,x_2),(y_1,y_2)\big) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_2y_1 \\
f_4\big((x_1,x_2),(y_1,y_2)\big) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_2y_2.
\end{aligned}$$

এদের চারজনের set-ই হবে একটা basis. একইভাবে $\mathbb{R}^n$-এর উপরে যাবতীয় bilinear form-এর vector space-এর একটা basis হবে $f_{ij}(\vec{x}, \vec{y}) = x_i y_j$ জাতীয় bilinear form-দের নিয়ে।

60.3.2 দুটো গুরুত্বপূর্ণ subspace

Theorem

Let V be any vector space over a field F. Let B be the vector space of all bilinear forms over V. Let S be the set of all symmetric bilinear forms in B. Then S is a subspace of B.

Proof: সাবধান, এই প্রমাণের জন্য কিন্তু matrix ব্যবহার করা যাবে না। কারণ V-কে finite dimensional বলে দেয় নি। অবশ্য প্রমাণটা তা বলে কঠিন কিছু নয়। একটা vector space-এর কোনো subset-কে একটা subspace দেখানোর কায়দা ছিল দুই ধাপে--এক, $\vec{0}$ তার মধ্যে আছে, আর দুই, এটা দেখানো যে, যদি subset-টার মধ্যে যা খুশী f, g নাও, এবং যেকোনো দুটো scalar নাও α, β, তবে $\alpha f + \beta g$ ফের subset-টার মধ্যেই থাকবে।

[6] মানে entry-গুলোর মধ্যে কোনো বিশেষ সম্পর্ক থাকা প্রয়োজন নেই।

> Clearly, the zero bilinear form, ie, $f(\vec{x},\vec{y}) \equiv 0$ is symmetric.

এবার দ্বিতীয় শর্তটা পরীক্ষা করা যাক--

> Let $f, g \in S$ and $\alpha, \beta \in F$.

আমরা দেখাতে চাই $\alpha f + \beta g \in S$ হবে, মানে যাই $\vec{x}, \vec{y}$ নাও না কেন, $(\alpha f + \beta g)(\vec{x},\vec{y}) = (\alpha f + \beta g)(\vec{y},\vec{x})$ হবে।

> Then
>
> $$\begin{aligned}(\alpha f + \beta g)(\vec{x},\vec{y}) &= \alpha f(\vec{x},\vec{y}) + \beta g(\vec{x},\vec{y}) \\ &= \alpha f(\vec{y},\vec{x}) + \beta g(\vec{y},\vec{x}) \quad [\because f, g \in S] \\ &= (\alpha f + \beta g)(\vec{y},\vec{x}).\end{aligned}$$
>
> Hence $(\alpha f + \beta g) \in S$.
> Thus S is a subspace of V, as required.

[Q.E.D]

Exercise 48: যদি উপরের theorem-টায় $\dim(V) = n$ হয়, তবে $\dim(S) = ?$ ■

একইরকম জিনিস করা যায় alternating bilinear form-দের নিয়েও--

THEOREM

Let V be any vector space over a field F. Let B be the vector space of all bilinear forms over V. Let A be the set of all alternating bilinear forms in B. Then A is a subspace of B.

Exercise 49: এই theorem-টা প্রমাণ করো। যদি $\dim(V) = n$ হয়, তবে $\dim(A) = ?$
HINT: আগের theorem-টার প্রমাণের অনুকরণে এগোও। Dimension-টাও বেরোবে $\dim(S)$ বার করার মত করেই।
■

S আর A দুজনেই হল B-এর দুটো subspace. এদের মধ্যে একটা সুন্দর সম্পর্ক আছে। নীচের অংকগুলো তোমাকে সেই সম্পর্কের দিকে ধাপে ধাপে এগিয়ে দেবে।

Exercise 50: Prove that every real bilinear form f can be written as $f = g + h$, where g is a symmetric bilinear form and h is an alternating bilinear form.
SOLUTION: কীভাবে g-টা নিলে সুবিধা হবে, ধরিয়ে দিচ্ছি--

$$g(\vec{x},\vec{y}) = \frac{1}{2}(f(\vec{x},\vec{y}) + f(\vec{y},\vec{x})).$$

■

Exercise 51: এমন একটা bilinear form দিতে পারো, যেটা একই সঙ্গে symmetric এবং alternating? ■

এই দুটো অংক থেকে যা শিখলে, সেটা মেলালেই পাবে নীচের theorem-টা--

> **THEOREM**
> If B is the vector space of all bilinear forms, and S and A are the subspaces formed by the symmetric and the alternating bilinear forms, then
> $$B = S \oplus A.$$

Exercise 52: Theorem-টা প্রমাণ করো। ■

সব শেষে একটা কথা বলি। **Bilinear transform** বলে অনেক সময়ে একটা কথা শুনবে। সেটার মানে কিন্তু অন্য। ওর মানে হল $f(z) = \frac{az+b}{cz+d}$ জাতীয় function, যেখানে $z \in \mathbb{C} \cup \{\infty\}$. এদের আরেকটা প্রচলিত নাম হল **Möbius transform**. আমাদের "বাংলায় বোঝানো ইংরাজী বই" সিরিজের Complex Analysis বইতে এদের বিষয়ে আলোচনা আছে।

Answers

1. **(1)** হ্যাঁ, **(2)** না, **(3)** না **(4)** হ্যাঁ **(5)** না।

2. **(1)** $\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. **(2)** $\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.
(3) $\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. **(4)** $\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 3 & 9 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

3. 30. **4.** না। একটা counterexample হল $f : \mathbb{R}^2 \to \mathbb{R}$ যেখানে $f(x,y) = \begin{cases} x^2 & \text{if } y = 0 \\ 2x^2 & \text{otherwise} \end{cases}$.

5. Symmetric matrix-দের linear combination-রাও symmetric matrix হয়। তাই vector space. একটা $n \times n$ সাইজের symmetric matrix-এর মধ্যে যে n^2-খানা entry আছে, তার মধ্যে খালি $\frac{n(n+1)}{2}$-খানাই free হয় (n-খানা diagonal entry, আর তার যেকোনো একপাশের entry-গুলো)। তাই dimension হল $\frac{n(n+1)}{2}$.

6. $\begin{bmatrix} 1.56 & 6 \\ 6 & 24 \end{bmatrix}$. **7.** $\frac{m}{2} \begin{bmatrix} a^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a^2 \end{bmatrix}$. **8.** $H(x,y,z) = \begin{bmatrix} e^x + 2z & -1 & 2x \\ -1 & 0 & 0 \\ 2x & 0 & 0 \end{bmatrix}$. সুতরাং $H(1,2,3) = \begin{bmatrix} e+6 & -1 & 2 \\ -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$. **9.** $\begin{bmatrix} 0 & 2 \\ 2 & 4 \end{bmatrix}$. **10.** Rank$= 5$, index$= 3$ এবং signature$= 1$.

11. **(1)** Nonsingular matrix দিয়ে গুণ করলে rank বদলায় না। আমাদের বেলায় $B = PAP'$, যেখানে P কোনো একটা nonsingular matrix. সুতরাং $r(A) = r(PA) = r(PAP') = r(B)$. **(2)** যেহেতু rank সমান, তাই diagonal বরাবর non-zero entry-দের সংখ্যা সমান। আবার যেহেতু A, B-এর সাইজ সমান, তাই diagonal বরাবর 0-র সংখ্যাও সমান হবে। **12.** না, কারণ symmetry বজায় রাখার জন্য column-দের নিয়েও একই কাজ করতে হত। ফলে (1,1)-th entry-র জায়গায় (2,2)-th entry-টা এসে উপস্থিত হত। কিন্তু সেটাও ছাই শূন্য! **13.**

The quadratic form is

$$\begin{bmatrix} x & y & z \end{bmatrix}\begin{bmatrix} 3 & -7 & 5 \\ -7 & 8 & 5 \\ 5 & 5 & -24 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

We apply congruent operations to the matrix to reduce it to a diagonal form.

$$\begin{bmatrix} 3 & -7 & 5 \\ -7 & 8 & 5 \\ 5 & 5 & -24 \end{bmatrix}$$

$$\xrightarrow[newR_3 = R_3 - \frac{5\,R_1}{3}]{newR_2 = R_2 + \frac{7\,R_1}{3}} \begin{bmatrix} 3 & -7 & 5 \\ 0 & -\mathbf{\frac{25}{3}} & \mathbf{\frac{50}{3}} \\ 0 & \mathbf{\frac{50}{3}} & -\mathbf{\frac{97}{3}} \end{bmatrix}$$

$$\xrightarrow[newC_3 = C_3 - \frac{5\,C_1}{3}]{newC_2 = C_2 + \frac{7\,C_1}{3}} \begin{bmatrix} 3 & 0 & 0 \\ 0 & -\frac{25}{3} & \frac{50}{3} \\ 0 & \frac{50}{3} & -\frac{97}{3} \end{bmatrix}$$

$$\xrightarrow{newR_3 = R_3 + 2\,R_2} \begin{bmatrix} 3 & 0 & 0 \\ 0 & -\frac{25}{3} & \frac{50}{3} \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{newC_3 = C_3 + 2\,C_2} \begin{bmatrix} 3 & 0 & 0 \\ 0 & -\frac{25}{3} & 0 \\ 0 & 0 & \mathbf{1} \end{bmatrix}$$

$$\xrightarrow{\substack{newR_1 = R_1/\sqrt{3} \\ newR_2 = R_2/\sqrt{\frac{25}{3}} \\ newC_1 = C_1/\sqrt{3} \\ newC_2 = C_2/\sqrt{\frac{25}{3}}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\substack{\text{Swap } R_2, R_3 \\ \text{Swap } C_2, C_3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Rank = no. of nonzero diagonals = 3.

Index = no. of positive diagonals = 2.

14. এই অংকটাও আগেরগুলোর মতই, জোড়ায় জোড়ায় congruent operation করে যাওয়া। প্রথমে ধরো row operation করলাম--

$$\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{3}{2} \\ \frac{1}{2} & -\frac{3}{2} & 0 \end{bmatrix}$$

$$\xrightarrow{newR_1 = R_2 + R_1} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & -\frac{3}{2} \\ \frac{1}{2} & -\frac{3}{2} & 0 \end{bmatrix}$$

তাহলে তার সাথে তাল মিলিয়ে column operation করতে হবে পরের ধাপেই--

$$\xrightarrow{newC_1 = C_2 + C_1} \begin{bmatrix} 1 & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & -\frac{3}{2} \\ -1 & -\frac{3}{2} & 0 \end{bmatrix}$$

এইভাবে চালিয়ে যাও--

$$\xrightarrow[newR_3 = R_3 + R_1]{newR_2 = R_2 - \frac{R_1}{2}} \begin{bmatrix} 1 & \frac{1}{2} & -1 \\ 0 & -\frac{1}{4} & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

$$\xrightarrow[newC_3 = C_3 + C_1]{newC_2 = C_2 - \frac{C_1}{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

$$\xrightarrow{newR_3 = R_3 - 4\,R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\xrightarrow{newC_3 = C_3 - 4\,C_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\begin{matrix} newR_2 = 2R_2 \\ newR_3 = R_3/\sqrt{3} \\ newC_2 = 2C_2 \\ newC_3 = C_3/\sqrt{3} \end{matrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\longrightarrow$$

$$\begin{matrix} \text{Swap } R_2, R_3 \\ \text{Swap } C_2, C_3 \end{matrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$\longrightarrow$$

Rank = no. of nonzero diagonals = 3

Index = no. of positive diagonals = 2

Signature = no. of positive diagonals - no. of negative diagonals = $2-1=-1$.

15. Canonical form: I_3. **16.** Normal form: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$, rank=3, signature=-1.

17. Normal form: $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$, rank=4, index=2, signature=0. **18.** (B). এখানে মূল matrix-টা হল $\begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}$ এবং তার normal form হল $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. **19.** (C). কারণ $5x^2+y^2+10z^2-4yz-10zx = 5(x^2-2zx+z^2)+(y^2-4yz+4z^2)+z^2 = 5(x-z)^2+(y-2z)^2+z^2$. তিনটে linear function-এর square-এর যোগফল। তাই signature হবে 3. **20.** $P = \begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{7} \\ 0 & 1 & \frac{1}{7} \\ 0 & 0 & 1 \end{bmatrix}$ নিলে $B = P'AP = diag\left(6, \frac{7}{3}, -\frac{26}{7}\right)$ আসা উচিত। মনে রেখো যে এটাই একমাত্র উত্তর নয়। যেমন ধরো যদি P-কে 2 দিয়ে গুণ করে দাও, তবেও $P'AP$-টা diagonal-ই থাকবে। **21.** $PAP' = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \frac{8}{3} \end{bmatrix}$, যেখানে $P = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1}{3} & -\frac{1}{3} & 1 \end{bmatrix}$. **22.**

22. $PAP' = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -\frac{4}{7} \end{bmatrix}$, যেখানে $P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -\frac{5}{7} & 0 & \frac{2}{7} \end{bmatrix}$. **23.** Zero matrix. **24.**

$$\begin{bmatrix} 6 & 0 & -6 \\ 0 & 1 & -2 \\ -6 & -2 & 18 \end{bmatrix}$$

$$\xrightarrow{newR_3 = R_3 + R_1} \begin{bmatrix} 6 & 0 & -6 \\ 0 & 1 & -2 \\ 0 & -2 & 12 \end{bmatrix}$$

$$\xrightarrow{newC_3 = C_3 + C_1} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 12 \end{bmatrix}$$

$$\xrightarrow{newR_3 = R_3 + 2\,R_2} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 8 \end{bmatrix}$$

$$\xrightarrow{newC_3 = C_3 + 2\,C_2} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

$$\xrightarrow{\substack{newR_1 = R_1/\sqrt{6} \\ newR_3 = R_3/\sqrt{8} \\ newC_1 = C_1/\sqrt{6} \\ newC_3 = C_3/\sqrt{8}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

All the diagonal entries of the normal form are strictly positive. So the quadratic form is positive definite.

25. এখানে প্রথম diagonal element-টা 0. সুতরাং প্রথম row-টা বিয়োগ করে বাকী row-গুলোর প্রথম element-দের শূন্য বানানো যাবে না। তাই আমরা প্রথমে row 1-এর সাথে row 2 যোগ করে নেব, যাতে diagonal element-য়া nonzero হয়। আমরা ইচ্ছে করলে row 1-এর সাথে row 3-ও যোগ করতে পারতাম, (1,1)-th element-টাকে nonzero বানানো নিয়ে কথা।

$$\begin{bmatrix} 0 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$\xrightarrow{newR_1 = R_2 + R_1} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 2 \\ \frac{1}{2} & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$\xrightarrow{newC_1 = C_2 + C_1} \begin{bmatrix} 1 & \frac{1}{2} & 2 \\ \frac{1}{2} & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$

$$\xrightarrow[newR_3 = R_3 - 2\,R_1]{newR_2 = R_2 - \frac{R_1}{2}} \begin{bmatrix} 1 & \frac{1}{2} & 2 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

$$\xrightarrow[newC_3 = C_3 - 2\,C_1]{newC_2 = C_2 - \frac{1}{2}C_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

$$\xrightarrow{newR_3 = R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

$$\xrightarrow{newC_3 = C_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

$$\begin{array}{c} newR_2 = 2R_2 \\ newR_3 = R_3/2 \\ newC_2 = 2C_2 \\ newC_3 = C_3/2 \\ \longrightarrow \end{array} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Rank = no. of nonzero diagonals = 3.

There is at least one positive and one negative diagonal entry. So the quadratic form is indefinite.

26. যদি A-টা nnd হয়, তবে তার normal form-টা হবে $P' \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} P$, যেখানে $r = r(A)$. এবার B নাও P-এর প্রথম r-খানা row দিয়ে তৈরী submatrix-টা। বিপরীতপক্ষে, $A = B'B$ হলে $\vec{x}'A\vec{x} = \|B\vec{x}\|^2 \geq 0$.
27. যদি A-টা nonsingular হয়, তবে A হবে p.d., সুতরাং আগের theorem থেকে পাই $|A| > 0$. যদি nonsingular না হয়, তবে $|A| = 0$. **28.** $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. **29.** **(1)** ঠিক, **(2)** ভুল, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$-র প্রথম row আর column নিয়ে তৈরী submatrix-টা হল $\begin{bmatrix} 1 \end{bmatrix}$, যেটা positive definite. **(3)** ঠিক। **(4)** ভুল, $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$-র প্রথম row আর column নিয়ে তৈরী submatrix-টা হল $\begin{bmatrix} 1 \end{bmatrix}$, যেটা positive definite.

30. ঠিক। **31.** (B). **32.** Hessian-টা সর্বত্রই $H = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. এদিকে $\nabla f(0,0) = (0,0)$. সুতরাং $(0,0)$ একটা stationary point. যেহেতু H হল 2×2 এবং $h_{11} > 0$ এবং $|H| > 0$, তাই H হল pd. সুতরাং local minimum. **33.** **(1)** Local minimum, কারণ $f(x,y) \geq 0$ আর $f(0,0) = 0$. **(2)** Saddle point, কারণ $f(0,0) = 0$ এবং $(0,0)$-র যত কাছেই যাও না কেন, এমন (x,y) পাবে যেখানে $f(x,y) > 0$, আবার এমন (x,y)-ও পাবে, যেখানে $f(x,y) < 0$. গুছিয়ে লিখলে $\forall \epsilon > 0 \quad \left(\frac{\epsilon}{2}, 0\right), \left(0, \frac{\epsilon}{2}\right) \in N_2\big((0,0), \epsilon\big)$ এবং $f\left(\frac{\epsilon}{2}, 0\right) > 0$ কিন্তু $f\left(0, \frac{\epsilon}{2}\right) < 0$. **34.** **(1)** না, **(2)** হ্যাঁ, **(3)** না, **(4)** হ্যাঁ, **(5)** না, **(6)** হ্যাঁ।

35. সরাসরি দেখাও যে, $f(a(x_1,y_1) + b(x_2,y_2), (w,z)) = af((x_1,y_1),(w,z)) + bf((x_2,y_2),(w,z))$, এবং $f((x,y), a(w_1,z_1) + b(w_2,z_2)) = af((x,y),(w_1,z_1)) + bf((x,y),(w_2,z_2))$.

36. $f\big((x_1,x_2),(y_1,y_2)\big) = f\big(x_1(1,0) + x_2(0,1), y_1(1,0) + y_2(0,1)\big) = x_1y_1f\big((1,0),(1,0)\big) + x_1y_2f\big((1,0),(0,1)\big) + x_2y_1f\big((0,1),(1,0)\big) + x_2y_2f\big((0,1),(0,1)\big) = \alpha x_1y_1 + \beta x_1y_2 + \gamma x_2y_1 + \delta x_2y_2$.

37. অসম্ভব। **38.** একটাই, শূন্য bilinear form-টা। **39.** **(1)** $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **(2)** $\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$.

40. $\frac{1}{2}(A + A')$. **41.** ভুল। যেমন $A = I_2$ আর $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ এবং $\vec{y} = -\vec{x}$. **42.** 0. কারণ একটাই bilinear form আছে, যেটা কখনও negative হয় না। সেটা হল শূন্য bilinear form. অন্য যেকোনো bilinear form-ই কোথাও না কোথাও negative হতে বাধ্য, কারণ যদি কোনো $\vec{a}, \vec{b}$-এর জন্য $f(\vec{a}, \vec{b}) \neq 0$ হলে $f(\vec{a}, \vec{b})$ আর $f(-\vec{a}, \vec{b})$-এর মধ্যে একজন তো negative হবেই! **43.** 0. কারণ এরকম খালি একটা bilinear form সম্ভব, সেটা হল শূন্য bilinear form. যদি $f(\vec{x}, \vec{y})$-টা শূন্য bilinear form না হয়, তবে কোথাও $f(\vec{a}, \vec{b}) \neq 0$ হবে। ধরো $|f(\vec{a}, \vec{b})| = \epsilon > 0$. তবে $K = \frac{11}{\epsilon}$ নাও। বুঝতেই পারছ $|f(K\vec{a}, \vec{b})| = 11 > 10$ হয়ে যাবে। **44.** **(1)** ঠিক, **(2)** ভুল।

46. **(1)** Symmetric,**(2)** symmetric,**(3)** alternating,**(4)** কোনোটাই নয়। **47.**

> Let $f : V \times V \to F$ be any alternating bilinear form. Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be any basis of V. Let $B \equiv ((b_{ij}))$ be the matrix of f wrt this basis. Then $\forall i, j \quad f(\vec{v}_i, \vec{v}_j) = b_{ij}$.
> Since f is alternating, $\forall i, j \quad f(\vec{v}_i, \vec{v}_j) = -f(\vec{v}_j, \vec{v}_i)$, ie, $b_{ij} = -b_{ji}$. Hence B is skew-symmetric, as required.

48. $\frac{n(n+1)}{2}$. কারণ S-এর প্রতিটা সদস্য মানে একটা $n \times n$ সাইজের symmetric matrix. এবার degrees of freedom দিয়ে চিন্তা করো। যেমন 2×2-এর বেলায় একটা symmetric matrix দেখতে $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$-এর মত। এর মধ্যে খালি তিনটে entry-ই free. তাই এক্ষেত্রে dimension হল 3. **49.** $\frac{n(n-1)}{2}$.

50. $h(\vec{x}, \vec{y}) = \frac{1}{2}(f(\vec{x}, \vec{y}) - f(\vec{y}, \vec{x}))$ নিলেই হবে। **51.** এরকম একটাই bilinear form আছে, সেটা হল শূন্য bilinear form. **52.** আমরা দেখিয়েছি S, A হল B-এর দুটো subspace. আরও দেখিয়েছি যে, $S = A + B$. এটাও দেখিয়েছি যে, $S \cap A = \{\vec{0}\}$, যেখানে $\vec{0}$ হল শূন্য bilinear form. সুতরাং $B = S \oplus A$.

Chapter VII
Eigenanalysis

এইবার আমরা এমন একটা জিনিস শিখব যার অপরিসীম গুরুত্ব, কিন্তু যেটা প্রথম প্রথম একেবারে অর্থহীন মনে হয়। অংকগুলো মোটেই শক্ত কিছু নয়, কিন্তু কেন কী হচ্ছে সেটা প্রথমে কিছুটা সময় নিয়ে বোঝা দরকার।

DAY 38
জিনিসটা কী?

38.1 গোড়ার কথা

একটা সহজ জিনিস দিয়ে শুরু করি। ধরো একটা matrix দিলাম--

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

যদি গ্রাফ কাগজে কোনো একটা বিন্দু নিই (x, y), তবে এই matrix-টা দিয়ে গুণ করে আরেকটা নতুন বিন্দু (X, Y) পেতে পারি এইভাবে--

$$\begin{bmatrix} X \\ Y \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2y \end{bmatrix}.$$

তার মানে $X = x$ আর $Y = 2y$. এঁকে দেখিয়েছি Fig 1-এ। ভাবতে পারো যেন (x, y) থেকে (X, Y) পাওয়া যায় vertical দিকে টেনে গ্রাফ কাগজটাকে দ্বিগুণ লম্বা করে দিলেই, কারণ x অপরিবর্তিত থাকছে, অথচ y হয়ে যাচ্ছে দ্বিগুণ। এবার ধরো গ্রাফের পাতায় একটা ছবি আঁকলাম Fig 2-এর মত। এর প্রতিটি বিন্দুকে যদি A দিয়ে গুণ করে দিই, তবে ছবিটা হয়ে দাঁড়াবে Fig 3-এর মত। অর্থাৎ বলা যেতে পারে যে, A হল vertical দিকে টেনে দ্বিগুণ করার matrix.

এইবার এই matrix-টা দেখো--

$$\begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix}.$$

Fig 1

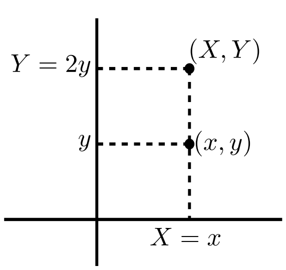


Fig 2

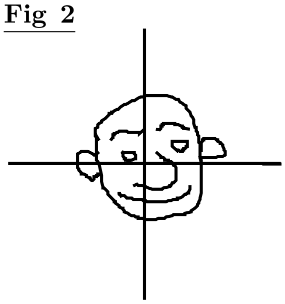
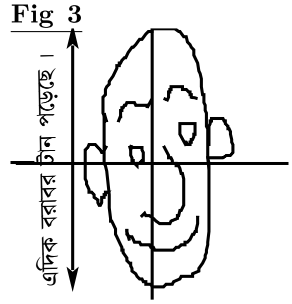


Fig 3



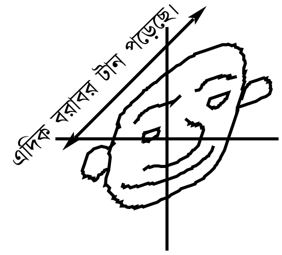


Fig 4

বেশ খটমট দেখতে, না? এবার আমরা এই matrix-টা দিয়ে Fig 2-এর প্রতিটা বিন্দুকে গুণ করব। ফলে ছবিটা দেখতে হয়েছে Fig 4-এর মত। এই ছবিটা আমরা কম্পিউটার দিয়ে বিন্দুগুলো বার করে করে এঁকেছি। এবার Fig 2-এর সঙ্গে Fig 4-এর তুলনা করলেই তুমি তৎক্ষণাৎ দেখবে যে, matrix-টা যতই খটমট দেখতে হোক, ওর প্রভাবে ছবিটার পরিবর্তনকে অতি সহজ ভাষায় প্রকাশ করা যায়-- ছবিটাকে খালি একটা হেলানো দিক বরাবর টেনে দ্বিগুণ করে দেওয়া হয়েছে। তার মানে একই পরিবর্তনকে আমরা দুইভাবে প্রকাশ করতে পারলাম--

- এক, খটমট matrix-টা দিয়ে,
- দুই, কোন্ দিকে টেনে দ্বিগুণ করতে হবে খালি সেটুকু বলে।

কোন্ বর্ণনাটা বেশী সহজ লাগল? নিশ্চয়ই দ্বিতীয়টা। প্রথম বর্ণনাটা এতই জটিল ছিল যে, আঁকতে কম্পিউটারের শরণাপন্ন হওয়া ছাড়া পথ ছিল না। অথচ দ্বিতীয় বর্ণনাটা জানা থাকলে হাতেই এঁকে ফেলা যেত। আরও একটা কারণেও দ্বিতীয় বর্ণনাটা বেশী সুবিধাজনক। Matrix দিয়ে করতে হলে পুরো কাজটা গ্রাফ কাগজে করতে হয়, কারণ প্রতিটি বিন্দুকে (x, y) আকারে প্রকাশ করতে হয়। কিন্তু দ্বিতীয় বর্ণনাটার ক্ষেত্রে সাদা কাগজেই করা যায়, খালি একটা লাইন টেনে দেখিয়ে দিতে হবে কোন্ দিক বরাবর টেনে দ্বিগুণ করতে হবে। সেটা গ্রাফ আঁকার থেকে অনেক সহজ। ঠিক যেমন, যদি তুমি জানতে চাও অমুক জায়গায় কী করে যেতে হবে, আমি তোমাকে সরাসরি এক দিকে আঙুল দেখিয়ে বলতে পারি "ওই দিকে যাও", কিংবা বলতে পারি "$(3, 5)$ দিক বরাবর যাও"। দ্বিতীয়ক্ষেত্রে তোমাকে জানতে হবে আমি কোন্ coordinate system-এর সাপেক্ষে কথাটা বলছি, কিন্তু সরাসরি আঙুল দিয়ে দেখিয়ে দিলে কোনো coordinate system-টিস্টেমের তোয়াক্কা করতে হয় না।

ঠিক একই কারণে "অমুক matrix-টা দিয়ে গুণ করো" বলার চেয়ে অনেক সময়েই "অমুক vector বরাবর ছবিটাকে টেনে এতটা লম্বা করে দাও" এই বর্ণনাটা বেশী স্বাভাবিক। এটা যেন ভিতরের কথাটাকে একবারে প্রকাশ করে দিল, কোনো coordinate system-এর প্রয়োজনই হল না। তাই এইরকম vector-কে আমরা বলি matrix-টার নিজস্ব vector. জার্মান ভাষায় "নিজস্ব"-কে বলে “eigen” (উচ্চারণ আইগেন, আইজেন নয়)। তাই জার্মান পণ্ডিতরা এই vector-টাকে বলেন matrix-টার একটা **eigenvector**. আর সেই vector বরাবর টেনে কতটা লম্বা করতে হবে, সেটাও matrix-টার নিজস্ব সম্পত্তি, একে বলে **eigenvalue**. ইংরাজিতে অনেক সময়ে এদের বলে **characteristic vector** আর **characteristic value**. কোনো matrix-কে তার eigenvalue আর eigenvector দিয়ে প্রকাশ করলে সুবিধা হয় বললাম। কিন্তু দুটো প্রশ্ন থেকেই যায়--

- এক, যে কোনো matrix-এর ক্ষেত্রেই কি এই রকম "কোনোদিকে টেনে লম্বা করা"-র ব্যাপারটা হয়?
- দুই, যদি কোনো matrix-এর বেলায় তা হয় বলে জানাও থাকে, তাহলেই বা কোনদিকে (মানে কোন eigenvector বরাবর) এবং কতটা (মানে কী eigenvalue পরিমাণ) টানতে হবে সেটা বুঝব? আমাদের উদাহরণে আমরা ছবিটা আগে কম্পিউটার দিয়ে এঁকে নিয়েছিলাম বলে বোঝা গেছিল। কম্পিউটার ছাড়া কি কাজটা করা যেত না?

এর মধ্যে দ্বিতীয় প্রশ্নটার উত্তর আগে দেব। ধরো আমরা গোড়ায় যে সহজ matrix-টা নিয়েছিলাম, সেটা নিয়েই শুরু করি--

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

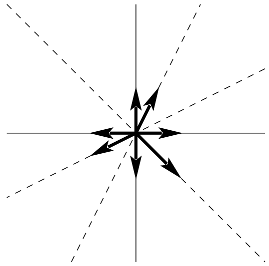


Fig 5

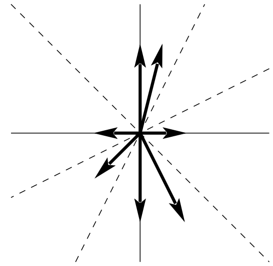


Fig 6

সেখানে আমরা অংক কষেই দেখে নিয়েছিলাম যে

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2y \end{bmatrix},$$

এবং তা থেকেই বলতে পেরেছিলাম যে, এটা vertical দিকে টেনে দ্বিগুণ লম্বা করে দেয়। এবার Fig 5-এর ছবিটা দেখো। এখানে খালি অনেকগুলো vector আঁকা আছে বিভিন্ন দিকে। যদি এই matrix-টা দিয়ে গুণ করি তবে ব্যাপারটা দাঁড়াবে Fig 6-এর মত। এখানে একটা জিনিস লক্ষ কর। যে vector-গুলো হেলানো ছিল, তারা গুণ করার ফলে নিজেদের লাইন থেকে সরে গেছে। কিন্তু vertical vector দুটো vertical-ই আছে। আবার horizontal vector দুটোও কিন্তু horizontal-ই আছে। এটা কেন হচ্ছে বোঝা কঠিন নয়, আমরা যেকোনো vector (x, y)-এর x অংশটাকে অপরিবর্তিত রেখে y-টাকে খালি দ্বিগুণ করছি। ফলে যাদের মধ্যে x, y দুজনেই $\neq 0$, সেগুলো দুই দিকে আলাদা টান খেয়ে খানিকটা ঘুরে যাচ্ছে। কিন্তু যাদের বেলায় x বা y একজন শূন্য, তারা খালি একদিকের টানটাই অনুভব করছে, ফলে খালি বড়জোর দৈর্ঘ্যে কিছু পরিবর্তন হচ্ছে, দিক বদলাচ্ছে না। সুতরাং আমরা এইভাবে এগোব--

> একটা square matrix হাতে পেলেই সেটা দিয়ে গুণ করে করে দেখব এমন কোনো vector $\vec{v}$ পাওয়া যাচ্ছে কিনা, যেটা গুণের পরেও একই লাইনে থাকছে, মানে $A\vec{v}$ আর $\vec{v}$ একই লাইনে থাকছে।

দুটো vector একই লাইনে থাকা মানে একটা আরেকটার multiple হওয়া। তার মানে আমরা এমন vector $\vec{v}$ খুঁজছি যাতে $A\vec{v}$ হয় $\vec{v}$-র একটা multiple, অর্থাৎ এমন λ থাকে যাতে

$$A\vec{v} = \lambda\vec{v}$$

হয়। তুমি হয়তো এটা দেখেই লাফিয়ে উঠে বলবে, এ তো পাওয়া যাবেই, $\vec{v} = \vec{0}$ আর λ যা খুশি নিলেই হল! ঠিক কথা, কিন্তু তাতে আমাদের কাজ হবে না। আমরা এমন একটা $\vec{v}$ চাই যেটা লাইন বদলাবে না, কিন্তু $\vec{0}$ হল একমাত্র vector যার আবার কোনো লাইনই হয় না! সুতরাং আমাদেরকে $\vec{v} \neq \vec{0}$ নিয়ে কাজ করতে হবে। অর্থাৎ--

> A যদি একটা square matrix হয়, তবে একটা $\vec{v} \neq \vec{0}$-কে তার **eigenvector** বলব যদি এমন একটা সংখ্যা λ থাকে, যাতে $A\vec{v} = \lambda\vec{v}$ হয়। এই অবস্থায় λ-কে বলব A-র একটা **eigenvalue.** এই λ কিন্তু শূন্য হতেই পারে।

এবার প্রশ্ন হল A দেওয়া থাকলে খালি অংক কষে কী করে $\vec{v}$ আর λ বার করব? এখানে তিনটে ব্যাপার লক্ষণীয়--

- এক, যদি $\vec{v}$-টা দেওয়া থাকত, তবে তার জন্য λ বার করাটা খুবই সহজ হত, $A\vec{v}$-টা $\vec{v}$-এর কতগুণ সেটা দেখে নিলেই চলত। অবশ্যই যেকোনো $\vec{v}$ হলে হবে না, $\vec{v}$-টাকে একটা eigenvector হতে হবে।
- দুই, যদি λ-টা আগেভাগে জেনে ফেলার কোনো উপায় থাকত, তবে $\vec{v}$-টা বার করে ফেলতাম একটা linear system of equations থেকে। কারণ $A\vec{v} = \lambda\vec{v}$, অর্থাৎ--

$$(\lambda I - A)\vec{v} = \vec{0}.$$

- তিন, আমরা এখনও এটাই জানি না যে, যেকোনো A-র জন্য এরকম $\vec{v}$ এবং λ আদৌ পাওয়া যায় কিনা!

সুতরাং সব মিলিয়ে বেশ ঘোরালো পরিস্থিতি। এবং মগজাস্ত্র প্রয়োগের শুভক্ষণ সমাগত! উপরের দুই নম্বর পয়েন্টটা দ্যাখো। আমরা এমন λ আর $\vec{v} \neq 0$ খুঁজছি, যাতে

$$(\lambda I - A)\vec{v} = \vec{0}$$

হয়। এদিকে বলেছে $\vec{v} \neq \vec{0}$, অথচ তাকে $(\lambda I - A)$ দিয়ে গুণ করলে $\vec{0}$ হয়ে যাচ্ছে। এমনটা খালি একভাবেই সম্ভব--যদি $(\lambda I - A)$ একটা singular matrix হয়। যদি এটা স্পষ্ট না হয়, তবে এইভাবে ভাবো--

> এখানে $\vec{v}$ একটা nonnull vector, সেটা রয়েছে $\mathcal{N}(\lambda I - A)$-এর মধ্যে। সুতরাং nullity$(\lambda I - A) > 0$. তাই বাধ্য হয়েই $r(\lambda I - A) < n$ হতে হবে (যেখানে A হল $n \times n$), কারণ rank আর nullity মিলে তো n হতে হবে!

এই ব্যাপারটাকে আমরা determinant দিয়ে লিখতে পারি--

$$|\lambda I - A| = 0.$$

এইটা খুব বড় সুখবর, কারণ এই শর্তটার মধ্যে কোনো $\vec{v}$ নেই, সুতরাং এইটা থেকে λ বার করতে পারার আশা আছে। আর সেটা হয়ে গেলে দুই নম্বর পয়েন্টের কায়দায় $\vec{v}$-ও বেরিয়ে যাবে।
দুটো উদাহরণ করে দেখে নিই ব্যাপারটা কী দাঁড়ালো।

Example 1: প্রথমে সেই সহজ উদাহরণটা--

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

এখানে

$$\lambda I - A = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \lambda - 1 & 0 \\ 0 & \lambda - 2 \end{bmatrix},$$

যার determinant হল

$$|\lambda I - A| = (\lambda - 1)(\lambda - 2).$$

এটা কখন শূন্য হবে? যখন $\lambda = 1$ বা 2 হবে। তার মানে দুটো eigenvalue পেলাম $1, 2$. এবার eigenvector বার করার পালা। প্রথমে $\lambda = 1$ বসিয়ে $\vec{v}$ বার করি--

$$(I - A)\vec{v} = \vec{0},$$

মানে

$$(I - A)\vec{v} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

সমাধান করলে পাই, $y = 0$, এবং x যা খুশি। তার মানে $(x, 0)$ আকারের সব vector-ই একটা করে eigenvector (অবশ্যই $x \neq 0$ নেব, কারণ আগেই বলেছি যে, $\vec{0}$-কে eigenvector ধরা হয় না)।
একইভাবে যদি $\lambda = 2$ বসাই, তবে $(2I - A)\vec{v} = \vec{0}$ সমাধান করতে হবে, মানে--

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

যা থেকে পাচ্ছি $x = 0$, আর y যা খুশি। সুতরাং $(0, y)$ আকারের যে কোনো vector-ই হল 2-এর জন্য একটা eigenvector, (অবশ্যই $\vec{0}$ এড়াবার জন্য $y \neq 0$ হতে হবে)। ■

এই অংক থেকে একটা জিনিস লক্ষ্য কর। প্রত্যেকটা eigenvalue-র জন্য কিন্তু অসংখ্য eigenvector পাওয়া গেল। তার কারণ একটা eigenvector পেলেই তার লাইন বরাবর সব nonzero vector-ই তো একেকটা eigenvector.

Example 2: এবার খটমট matrix-টায় হাত দিই। এখানে

$$|\lambda I - A| = \begin{vmatrix} \lambda - \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \lambda - \frac{3}{2} \end{vmatrix} = \lambda^2 - 3\lambda + 2.$$

সমাধান করে এখানেও λ-র দুটো value পাওয়া যাচ্ছে, 1 আর 2. করেই দেখো! যদি $\lambda = 1$ বসাই, তবে $(\lambda I - A)\vec{v} = \vec{0}$ থেকে পাব--

$$\begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

যা থেকে আসবে $x = -y$. সুতরাং $(x, -x)$ চেহারার যেকোনো vector-ই হবে $\lambda = 1$-এর জন্য একেকটা eigenvector (খালি $x \neq 0$ হতে হবে)।
যদি $\lambda = 2$ বসাও তবে দেখবে যে, eigenvector-রা হবে (x, x) চেহারার vector-রা $(x \neq 0)$. আসলে eigenvector-রা তো একটা দিক নির্দেশ করে, তাই সেই দিক বরাবর যাবতীয় nonzero vector-ই একেকটা eigenvector. ■

আচ্ছা, সব সময়েই কি এই দুটো উদাহরণের মত এগোনো যাবে? যদি একবার λ পেয়ে যাই, তবে তার জন্য $\vec{v}$ সব সময়েই পাওয়া যাবে, কিন্তু $|\lambda I - A| = 0$ সব সময়ে সমাধান করা যাবে কিনা বোঝা যাচ্ছে না। যেমন ধরো নীচের অংকটায়।

Example 3: আগের উদাহরণ দুটোয় যা করেছিলাম, সেটা

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

নিয়ে কর।
SOLUTION: এবার $|\lambda I - A| = \lambda^2 + 1 = 0$ কী করে সমাধান করবে? তার জন্য complex number ব্যবহার করা ছাড়া পথ নেই। করলে পাব $\lambda = \pm i$. কিন্তু ছবি দিয়ে এর মানে কী হবে? $\lambda = 2$ হলে বলা চলত টেনে দ্বিগুণ করা, কিন্তু complex সংখ্যার ক্ষেত্রে এমন কিছু বলা শক্ত। আমি অন্ততঃ এর কোনো ভদ্র ব্যাখ্যা জানিনা, বা অন্য কেউ জানে বলেও শুনি নি। ■

যারা complex সংখ্যা ভলোবাসো না, তারা আমার উপর চটে ওঠার আগে মনে রেখো যে, আমি কিন্তু শখ করে complex সংখ্যাদের ডেকে আনিনি, ওরা নিজেরাই এসে উপস্থিত হয়েছে, কারণ যদি field-টা $\mathbb{R}$-ও নিই, তবে $det(\lambda I - A)$ সর্বদাই একটা real polynomial হবে বটে, কিন্তু তার সব root-ই যে real number হবে, এমন তো কোনো কথা নেই! যদি polynomial-টার degree বড় কিছু হয়, তবে তার real root আদৌ থাকবে কি না, বা থাকলেও কটা থাকবে, সেটা জানারও কোনো সহজ পথ নেই। কিন্তু যদি field-টাকে $\mathbb{C}$ নিই, তবে দিব্যি একটা theorem আছে যে, একটা complex polynomial[1]-এর যত degree ঠিক ততগুলো complex root পাবই। আর real root-রা তো এর ফলে বাদ পড়ে যাচ্ছে না, ওরাও তো একধরণের complex root, যেমন $\lambda = 3$ যদি একটা root হয়, ওটাকে $\lambda = 3 + 0i$ ভাবলেই হল! এই কারণেই আমরা eigenvalue, eigenvector-এর আলোচনা করার সময়ে সর্বদা field-টাকে $\mathbb{C}$ ধরি। তবে প্রয়োগের সময়ে অধিকাংশ ক্ষেত্রেই খালি real eigenvalue-দের নিয়েই মাথা ঘামাই।
লম্বা আলোচনা হল, এইবার সব কিছু অংকের ভাষায় গুছিয়ে ফেলার সময় এসেছে।

38.2 অংকের ভাষায়

Example 4: Define eigenvalue and eigenvector of a square matrix.[1] **(2019)**

SOLUTION:

[1]Complex polynomial মানে এমন একটা polynomial, যার coefficient-গুলো একেকটা complex সংখ্যা।

DEFINITION: Eigenvalue, eigenvector

Let A be a square matrix over $\mathbb{C}$. Let $\vec{v} \in \mathbb{C}^n$ be a nonnull vector and $\lambda \in \mathbb{C}$ be such that

$$A\vec{v} = \lambda\vec{v}.$$

Then we shall call λ an **eigenvalue** of A and $\vec{v}$ a corresponding **eigenvector** of A.

■

Example 5: Show that 2 is an eigenvalue of the matrix

$$\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

and also find the eigenvectors corresponding to this eigenvalue.
SOLUTION:

Let the given matrix be A. Shall show that $\exists \vec{v} \neq \vec{0}$ $\;A\vec{v} = 2\vec{v}$, ie, $(A-2I)\vec{v} = \vec{0}$.

Now

$$A - 2I = \begin{bmatrix} 6-2 & -2 & 2 \\ -2 & 3-2 & -1 \\ 2 & -1 & 3-2 \end{bmatrix} = \begin{bmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix}.$$

Solving the system

$$\begin{bmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

we have $2x - y + z = 0$.

যেহেতু এই শর্ত পালন করে অন্ততঃ একটা nonnull vector আছে, তাই 2 সত্যিই একটা eigenvalue. এরকম একটা nonnull vector-এর উদাহরণ লিখেই দিই--

In particular, taking $x = 1,\ y = 2$ and $z = 0$ we get $\vec{v} = (1, 2, 0) \neq \vec{0}$. So 2 is an eigenvalue of A, as required.

Also, the set of all eigenvectors of A corresponding to the eigenvalue 2 is

$$\{(x, y, z) \neq (0, 0, 0) \;\; : \;\; 2x - y + z = 0\}.$$

■

আগের অংকে matrix-টা দেওয়া ছিল, সেখান থেকে আমরা eigenvalue আর eigenvector-এর আলোচনায় গেছিলাম। এবার উল্টোদিকটা দেখব, যেখানে eigenvalue আর eigenvector-গুলো দেওয়া থাকবে। তা থেকে matrix-টা বানাতে হবে।

Example 6: Let A be a 3×3 real matrix having the eigenvalues $1, 2, 0$. Let $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}$ be the eigenvectors of A corresponding to the eigenvalues $1, 2, 0$, respectively. Find the matrix A.[5] **(2020.CC6.3g)**

SOLUTION: প্রথমে vector তিনটের নাম দিয়ে নিই--

> Let the given eigenvectors be called $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$.

এই তিনটে vector-এর একটা বৈশিষ্ট্য আছে, যেটা ব্যবহার করে আমরা অংকটা করব। বৈশিষ্ট্যটা হল--এরা মিলে একটা basis হয়।

> Then $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a basis of $\mathbb{C}^3$.

আমরা জানি A দিয়ে গুণ করলে এই তিনটে vector কোথায় যায়--

> Also $A\vec{v}_1 = \vec{v}_1,\ A\vec{v}_2 = 2\vec{v}_2$ and $A\vec{v}_3 = \vec{0}$.

সুতরাং অন্য যেকোনো vector দেওয়া থাকলে, সেটাকেও এই basis-এর সাপেক্ষে প্রকাশ করে নেব। তাহলে বার করে ফেলতে পারব যে, সেই vector-টাকে A দিয়ে গুণ করলে কোথায় যাবে--

> Now any $\vec{v} \equiv (x, y, z) \in \mathbb{C}^3$ may be expressed as $a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3$ for some $a_1, a_2, a_3 \in \mathbb{C}$.
> Then
> $$\begin{bmatrix}x\\y\\z\end{bmatrix} = a_1\begin{bmatrix}1\\2\\1\end{bmatrix} + a_2\begin{bmatrix}0\\1\\1\end{bmatrix} + a_3\begin{bmatrix}1\\1\\1\end{bmatrix}.$$
> Solving we get $a_1 = y - z,\ a_2 = z - x,\ a_3 = z - y + x$.

এখানে সমাধানটা রাফে করে দেখে নিও, ঠিক বলছি কিনা।

Hence

$$\begin{aligned} A\vec{v} &= a_1A\vec{v}_1 + a_2A\vec{v}_2 + a_3A\vec{v}_3 \\ &= a_1\vec{v}_1 + 2a_2\vec{v}_2 + \vec{0} \\ &= (y-z)\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + 2(z-x)\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} y-z \\ 2y-2x \\ y+z-2x \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & -1 \\ -2 & 2 & 0 \\ -2 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}. \end{aligned}$$

Since this holds for all $(x, y, z) \in \mathbb{C}^3$, we have

$$A = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 2 & 0 \\ -2 & 1 & 1 \end{bmatrix}.$$

■

এইসব অংকে eigenvalue-গুলো দেওয়া থাকছিল। যদি না দেওয়া থাকত, তবে আমরা দেখেছি কী করে eigenvalue বার করা যায় $|\lambda I - A| = 0$ সমাধান করে। এরও একটা আলাদা নাম আছে--

DEFINITION: Characteristic polynomial/equation

Let $A_{n\times n}$ be a square matrix over $\mathbb{C}$. Then its **characteristic polynomial** is defined as $p(\lambda) = |\lambda I - A|$, which is a monic polynomial of degree n in λ. The equation $|\lambda I - A| = 0$ is called the **characteristic equation** of A.

এখানে দুটো জিনিস বলা দরকার--

- এক, **monic polynomial** কাকে বলে? এর মানে এমন একটা polynomial, যার সবচেয়ে বেশী power-ওয়ালা term-টার coefficient হল 1. যেমন $\lambda^2 - 3\lambda + 4$ হল monic, কিন্তু $-\lambda^2 - 3\lambda + 4$ বা $2\lambda^2 + \lambda + 1$ মোটেই monic নয়।
- কোনো কোনো বইতে দেখবে characteristic polynomial-এর সংজ্ঞা দেওয়া থাকে $|A - \lambda I|$. এই সংজ্ঞা নিলে characteristic polynomial-টা monic নাও থাকতে পারে। অবশ্য characteristic equation-এর সংজ্ঞা দুই ক্ষেত্রেই একই থাকে।

Example 7: যদি $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ হয়, তবে তার characteristic polynomial বার করো।

SOLUTION: উত্তর হল--

$$|\lambda I - A| = \left| \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right| = \begin{vmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 4 \end{vmatrix} = (\lambda - 1)(\lambda - 4) - 6 = \lambda^2 - 5\lambda - 6.$$

লক্ষ্য করো কেন এটা একটা 2 degree-র polynomial হল। $\lambda I - A$-র মধ্যে diagonal-গুলো "$\lambda-$ কিছু একটা" আকারের, বাকি entry-গুলো সংখ্যা। এরকম জিনিসদের গুণ করলে polynomial ছাড়া হবেটা কী? একটা $n \times n$ সাইজের matrix-এর determinant বার করার জন্য যে $n!$-খানা term থাকে, তার প্রত্যেকটাই n-খানা entry গুণ করে পাওয়া, তাই কোনো term-এর degree-ই n-এর বেশী হতে পারে না। Degree-টা n হতে পারে একমাত্র একটা term-এর বেলাতেই, যেটা সবগুলো diagonal entry গুণ করে পাওয়া গেছে। সেখানে λ^n-এর coefficient হল 1. সেই জন্যই monic হচ্ছে। ■

Exercise 1: Find the characteristic polynomial of the following matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix}.$$

■

এবার তাহলে eigenvalue বার করার মূলমন্ত্রটা একটা theorem-এর আকারে লিখে ফেলি--

THEOREM

Let A be a square matrix with complex entries. Then $\lambda \in \mathbb{C}$ is an eigenvalue of A iff it is a root of the characteristic equation, *i.e.*, $|\lambda I - A| = 0$.

Characteristic equation-এর root হিসেবে বার করা যায় বলে eigenvalue-র আরেক নাম **characteristic root**.

Exercise 2: Find the eigenvalues and corresponding eigenvectors of $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.[3] **(2020.CC6.3cii)**

HINT:

এখানে diagonal matrix নিয়ে কাজ হচ্ছে, তাই determinant বার করা সহজ, diagonal entry-গুলোকে গুণ করে দিলেই হল।

Here the characteristic polynomial is

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 0 & 0 \\ 0 & \lambda - 3 & 0 \\ 0 & 0 & \lambda - 5 \end{vmatrix} = (\lambda - 2)(\lambda - 3)(\lambda - 5).$$

So the roots of the characteristic equation are 2, 3 and 5.

For $\lambda = 2$, a vector $\vec{v} \neq \vec{0}$ is an eigenvector iff $(2I - A)\vec{v} = \vec{0}$, ie,

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

> Solving, we get x arbitrary, $y = z = 0$. So the set of all eigenvectors for eigenvalue 2 is $\{(x, 0, 0) \ : \ x \neq 0\}$.

বাকী দুটো তোমার জন্য রেখে দিলাম। ■

একটা $\vec{v}$ যদি কোনো eigenvalue-এর জন্য eigenvector হয়, তবে তার সব nonzero multiple-ও যে ওই একই eigenvalue-এর জন্য eigenvector হবে, সে তো বলেইছি। একইরকম আরেকটা জিনিস বলব এবার। ধরো $\vec{u}, \vec{v}$ দুজনেই একই eigenvalue-র জন্য দুটো eigenvector. মানে eigenvalue-টাকে λ নাম দিলে $A\vec{u} = \lambda\vec{u}$ এবং $A\vec{v} = \lambda\vec{v}$. এবার যদি ওদের যা খুশি একটা linear combination নিই, ধরো $\vec{w} = a\vec{u} + b\vec{v}$, তবে বুঝতেই পারছ যে, $A\vec{w} = \lambda\vec{w}$ হবে, কারণ--

$$A\vec{w} = A(a\vec{u} + b\vec{v}) = aA\vec{u} + bA\vec{v} = a\lambda\vec{u} + b\lambda\vec{v} = \lambda\vec{w}.$$

তার মানে একই eigenvalue-এর জন্য যাবতীয় eigenvector-দের set-টা হল closed under linear combination, সুতরাং একটা subspace. এখানে অবশ্য একটা ছোট্টো সমস্যা আছে। Subspace-এর মধ্যে $\vec{0}$ থাকবেই, অথচ সেটা তো আবার একটা eigenvector হতে পারে না। সুতরাং ঠিক করে বললে বলা উচিত যে, একই eigenvalue-এর জন্য যাবতীয় eigenvector-রা, এবং $\vec{0}$ মিলে যে set-টা হয়, সেটা হল একটা subspace. এর নাম হল ওই eigenvalue-এর জন্য **eigenspace.** লেখার সময়ে অবশ্য এটাকে বেশ সংক্ষেপে লেখা যায়--

DEFINITION: Eigenspace

Let $A_{n\times n}$ have an eigenvalue λ. Then $\mathscr{N}(\lambda I - A)$ is called the **eigenspace** for λ. This is a subspace of $\mathbb{C}^n$.

এখানে nullspace-এর গল্প কী করে এল, বুঝলে তো? যেসব $\vec{v}$-এর জন্য $(\lambda I - A)\vec{v} = \vec{0}$ হয়, তাদের set-টাকেই তো বলে $\mathscr{N}(\lambda I - A)$. এর মধ্যে $\vec{0}$-ও আছে।

Example 8: আগের অংকে $\lambda = 2$-এর জন্য eigenspace কী ছিল?

SOLUTION: পুরো x-axis-টা। ■

Exercise 3: Obtain the eigenvalues, eigenvectors and eigenspaces of the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

[5] **(2021.CC12.3e)** ■

এতক্ষণ যে কটা অংক দেখলাম, প্রত্যেকটাতেই determinant বার করা সহজ হচ্ছিল। সেটা সব সময়ে নাও হতে পারে। সেক্ষেত্রে কষ্ট করেই determinant বার করতে হবে। যেমন নীচের অংকটায়।

Exercise 4: The eigenvalues of $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$ are

(A) $1, 2, 2$ (B) $1, -1, 2$ (C) $3, 3, 8$ (D) $3, 6, 8$.

(2022.CC6.1j)

HINT:

লক্ষ্য করো, অনেক সময়ে একই সংখ্যা একাধিক বার লেখা হয়েছে, যেমন প্রথম option-এ 2 লেখা হয়েছে দুবার। এর মানে characteristic polynomial-এ $\lambda - 2$ দুবার এসেছে, অর্থাৎ $(\lambda - 2)^2$ হিসেবে। সুতরাং প্রথম option-টা বলছে যে, characteristic polynomial-টা হল $(\lambda - 1)(\lambda - 2)^2$. ■

DAY 39
খায় না মাথায় মাখে?

ছাত্রছাত্রীদের মুখে এরকম অভিযোগ প্রায়ই শুনি--অংক তো চোখকান বুঁজে তো করে আসছি বহুদিন যাবৎ, কিন্তু কেন যে কী করছি, সেটা নিজেই জানি না। যদি প্রশ্ন করি অংকের কোন জিনিসটা নিয়ে এই আপত্তিটা সবচেয়ে বেশী, তবে অনেকের কাছেই উত্তর শুনেছি--eigenvalue আর eigenvector! তাই এখানে আমরা গোড়াতেই তিনটে অত্যন্ত গুরুত্বপূর্ণ প্রয়োগের কথা জেনে রাখব। প্রতিটা প্রয়োগের মধ্যেই অংকের খুঁটিনাটি আছে, কিন্তু তা নিয়ে আমরা আপাততঃ মাথা ঘামাব না। প্রমাণ ইত্যাদি যথাসময়ে আসবে, এখন আমাদের লক্ষ্য কেবল মূল ধারণাটা বুঝে রাখা।

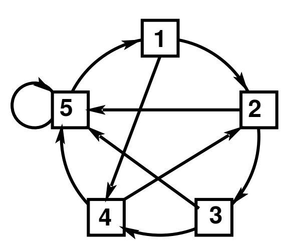
39.1 Google page rank

প্রথম প্রয়োগটা একেবারে অত্যাধুনিক, এবং বিশ্বে তার সুদূরপ্রসারী প্রভাব। তোমরা যারা ইন্টারনেট কর, তারা সকলেই Google-এর নাম শুনেছ। তুমি যদি কোনো জিনিস ইন্টারনেটে খুঁজে পেতে চাও, তার কথা Google-কে বলে দিলেই চোখের নিমেষে সেই বিষয়ে কোথায় কোন website আছে, সব চলে আসে। বোঝাই যায় যে, Google যাবতীয় website-এর একটা তালিকা কোথাও বানিয়ে রেখেছে। সেটা একটা বিশাল তালিকা সন্দেহ নেই, এবং সেটা আগাগোড়া খুঁজে ফেলাও সহজ কথা নয়, কিন্তু Google-এর আসল কেরামতি সেখানেই নয়। Google কোনো এক আশ্চর্য কায়দায় বুঝে ফেলে যে, কোন website-গুলো বেশী দরকারী, এবং সেইভাবে website-গুলোকে সাজিয়ে তোমার সামনে এনে ফেলে। এই কায়দাটার আবিষ্কর্তা হলেন Larry Page, যিনি Google-এর দুই প্রতিষ্ঠাতার একজন। কায়দাটার নাম রাখা হয়েছে Page Ranking, এখানে page শব্দটায় webpage-ও বোঝানো হল, আবার Larry Page-ও বোঝানো হল। শুনলে অবাক হবে যে, এর মূলে[2] আছে eigenvector বার করা। একটা উদাহরণ নিলে বুঝতে সুবিধা হবে।

Example 9: Fig 7-এ একটা ছোট্টো ইন্টারনেট এঁকেছি। এখানে খালি 5-টা webpage আছে, যাদের [1], [2], [3],... ইত্যাদি নাম দিয়েছি। যেমন ধরো [1] হয়তো www.facebook.com, আবার [2] হয়তো yahoo.co.in, এইরকম। প্রত্যেকটার থেকে দুটো করে link বেরিয়েছে, যেগুলোকে তীর চিহ্ন এঁকে দেখিয়েছি। [5] থেকে একটা link আবার [5]-এই ফিরে এসেছে। Fig 7-এর দিকে তাকিয়ে বলতে পারো, কোন webpage-টাকে সবচেয়ে গুরুত্বপূর্ণ মনে হচ্ছে?

[2] অবশ্য কে Google-কে বেশী টাকা দেয়, বা তুমি কোন্ দেশ থেকে সার্চ করছ, এগুলোও হিসেবে নেওয়া হয়।

Fig 7



SOLUTION: নিশ্চয়ই $\boxed{5}$-কে, কারণ অধিকাংশ page-ই $\boxed{5}$-এ link দিয়ে রেখেছে। এখানে আমরা একটা page-এর গুরুত্বের মাপকাঠি হিসেবে নিলাম সেই page-এ কতগুলো link এসে পৌঁছেছে, সেটা। কিন্তু এটা খুব ভালো মাপকাঠি নয়। যেমন ধরো $\boxed{1}$ এবং $\boxed{3}$ দুটোতেই ঠিক একটা করে link এসেছে। তা বলে তুমি কি বলবে যে ওরা সমান গুরুত্বপূর্ণ? না, কারণ $\boxed{1}$-এর বেলায় link-টা এসেছে সেই ভীষণ জনপ্রিয় $\boxed{5}$ নম্বর page থেকে, আর $\boxed{3}$-এর বেলায় link-টা এসেছে অখ্যাত $\boxed{2}$ থেকে। হাজার হোক আমি আমার page থেকে তোমাকে একটা link দিলে তোমার যত জনপ্রিয়তা বাড়বে, তার চেয়ে সেই link-টাই যদি আসত Times of India-র page থেকে, তবে তার কৌলীন্যই আলাদা হত। সুতরাং কতগুলো link আসছে সেটা একেবারে ফেলনা না হলেও, কেবলমাত্র সেই সংখ্যাটুকু দিয়েই খালি কোনো page-এর গুরুত্ব মাপা উচিত নয়। গুরুত্ব মাপার জন্য আরো সূক্ষ্মভাবে এগোতে হবে। এবং সেখানেই Google page rank-এর বাহাদুরি। ■

এবার জিনিসটা কী সেটা বলি। প্রথমে একটা সহজ ধারণা বোঝা দরকার।

Example 10: ধরো একজন লোক প্রথম কলকাতায় এসেছে। সে কিছুই চেনে না। সে যদি কেবল উদ্দেশ্যহীনভাবে কলকাতার পথে পথে যেখানে যা বাস পায় তাতেই উঠে পড়ে, এবং বাসের কণ্ডাক্টরদের চিৎকারগুলো শোনে, তবে কেবল তা থেকেই সে মালুম করতে পারবে যে, কলকাতায় "ধর্মতলা" এবং "ডালস্যি"[3] নামে দুটো জায়গা আছে। এবং যেহেতু এত বাসরুট ওখানে যাচ্ছে, কলকাতার জীবনযাত্রায় ওদের গুরুত্ব নিয়ে তার মনে সন্দেহের অবকাশ থাকবে না। ■

ঠিক এই ধারণাটাই Google কাজে লাগায়। ধরো লোকটা কলকাতার পথে পথে যা খুশি বাসে উঠে পড়ার বদলে একই কাজ করছে ইন্টারনেটে। সে random-ভাবে browse করে চলেছে। একটা page-এ এসে random কোনো একটা link-এ কিছু মাত্র না ভেবেই দুম করে click করে দিচ্ছে। এইভাবে সে এক page থেকে আরেক page-এ যাচ্ছে, মাঝে মাঝে আবার পুরোনো কোনো page-এ ফিরেও আসছে। এইভাবে চললে যে page-গুলোতে সে বার বার করে আসবে, সেগুলোই নিশ্চয়ই গুরুত্বপূর্ণ page হবে। সুতরাং কোথায় সে কতবার আসছে, সেটা দিয়েই সেই page-এর জনপ্রিয়তা মাপা যাবে।
সমস্যা খালি একটাই--ইন্টারনেট তো আর একটুখানি জিনিস নয়, তার মধ্যে page আর link-এর সংখ্যা সু-বি-শা-ল! সুতরাং এইরকম উদ্দেশ্যবিহীনভাবে click করে করে পুরো ইন্টারনেটের খবর নিতে গেলে সে কাজ এ জন্মে শেষ হবে না।
এবং এখানেই মগজাস্ত্র! সত্যি সত্যি এরকমভাবে browse না করেই কেবলমাত্র অংক কষে বলে দেওয়া যায় উত্তরটা কী হবে, eigenvector ব্যবহার করে!!! এবং Google-এর page ranking ঠিক সেটাই করে!
ব্যাপারটা বোঝার জন্য আবার আমাদের ছোট্টো ইন্টারনেটে ফিরে আসি। ছোট্টো ইন্টারনেটের সুবিধা এই যে, তুমি সত্যি সত্যি এরকম উদ্দেশ্যবিহীনভাবে browse করে দেখতে পারো। নীচের উদাহরণে এই কাজটা হাতেকলমে করে দেখলে মজা পাবে।

Example 11: ধরো তুমি browse করা শুরু করলে $\boxed{1}$ থেকে। একটা কয়েন নাও। কয়েনটাকে টস্ কর। যদি head পড়ে, তবে চক্রাকারে ঘড়ির কাঁটার মত যে তীরটা আছে, সেটা ধরে এগোবে, আর tail পড়লে অন্য তীরটা নেবে, মানে $\boxed{4}$-এ চলে যাবে। আবার কয়েনটা টস্ কর, সেই একই নিয়ম, head পড়লে চক্রাকারে এগোনো, tail পড়লে অন্য তীর চিহ্নটা নেবে। এইভাবে ধরো 30 বার করলে। বুঝতেই পারছ যে, এটা আসলে একটা সাপ-লুডো খেলার মত। খালি ছক্কার জায়গায় কয়েন আর সাপ বা মইয়ের বদলে তীর চিহ্ন। প্রত্যেক ধাপে লিখে রেখো তুমি কোথায় আছো। যেমন ধরো যদি প্রথম তিনটে টসে head, head, tail আসে, তবে তুমি এগোবে এইভাবে--

$$\boxed{1} \xrightarrow{\quad head \quad} \boxed{2} \xrightarrow{\quad head \quad} \boxed{3} \xrightarrow{\quad tail \quad} \boxed{5}.$$

তারপর বার কর কোনখানে কতবার ছিলে। আমার পক্ষে অবশ্যই জানা সম্ভব নয় তোমার কয়েনে কখন কী পড়েছে। কিন্তু আমি লিখে দিচ্ছি যে, তুমি $\boxed{1}$-এ কাটিয়েছ মোটামুটি 8-বার, $\boxed{2}$-তে মোটামুটি 4 বার, $\boxed{3}$-এ 2 বার, $\boxed{4}$-এ 1 বার, এবং $\boxed{5}$-এ 15 বার। একটু উনিশ-বিশ হতে পারে, কিন্তু মোটামুটি এরকমই একটা কিছু হবেই! হাতে সময় (এবং আঙুলে জোর) থাকলে আরো বেশী বার টস্ করে দেখতে পারো। সবক্ষেত্রেই দেখবে মোটামুটিভাবে বিভিন্ন page-এ তুমি সময় কাটাচ্ছ এইরকম হারে--

---

[3]কলকাতা-অনভিজ্ঞ পাঠকের সুবিধার্থে জানিয়ে রাখি যে, ডালস্যির সঙ্গে ডাল বা লস্যির কোনো সম্পর্ক নেই। ওটা Dalhousie Square নামক জায়গার কণ্ডাক্টরি অপভ্রংশ।

1	2	3	4	5
$\frac{1}{4}$	$\frac{1}{7}$	$\frac{1}{14}$	$\frac{1}{28}$	$\frac{1}{2}$

যত বেশীবার টস্ করবে, ততই তুমি এদের বেশী কাছে এগিয়ে যাবে। ভাবতে গেলে এটা খুবই অদ্ভুত, কারণ যত বেশী টস্ করছ, ততই তো ব্যাপারটা আরও এলোমেলো হয়ে যাওয়া উচিত! কিন্তু আসলে তা হবে না (সেটাও অংক কষে প্রমাণ করা যায়)। এই সংখ্যাগুলোই হল Google-এর দেওয়া page rank. যেমন, এখানে সবচেয়ে বেশী page rank পেয়েছে $\boxed{5}$, তারপর $\boxed{1}$, ইত্যাদি। ■

এইটা হল হাতেকলমে করা, যেটা সত্যিকারের ইন্টারনেটের বেলায় সম্ভব নয়। এবার দেখি eigenvector ব্যবহার করে কীভাবে কাজটা সহজে করা যায়।
আমরা বার করতে চাইছি, random-ভাবে browse করলে আমরা বিভিন্ন page-এ কী হারে ফিরে ফিরে আসব। ধরো $\boxed{1}$ নম্বর page-এর বেলায় এই হারটা হল p_1, মানে যদি লুডোটা 1000 বার খেলি, তবে মোটামুটিভাবে $1000p_1$ বার আসব $\boxed{1}$ নম্বর page-এ। একইভাবে বাকিদের বেলায় $p_2, ..., p_5$ নাম দিই। এবার ধরো তোমাকে p_1 আর p_4 বলে দিলাম, তা থেকে কি p_2 বার করা সম্ভব? উত্তর হল--হ্যাঁ। তুমি যখনই $\boxed{1}$-এ আসবে, তার পর 50% সময়ে $\boxed{2}$-তে যাবে। আবার $\boxed{4}$ থেকেও 50% সময়ে $\boxed{2}$-তে হাজির হবে। যেহেতু এই দুই পথ ছাড়া $\boxed{2}$-তে আসার আর কোনো পথ নেই, তাই অবশ্যই--

$$p_2 = p_1 \times \tfrac{1}{2} + p_4 \times \tfrac{1}{2}.$$

যুক্তিটা খুব ভালো করে ভেবে দেখো। সেই একই যুক্তি থেকে এটাও আসবে--

$$p_3 = p_2 \times \tfrac{1}{2},$$

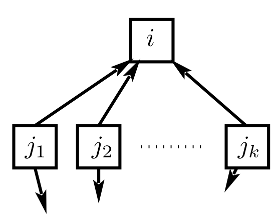
কারণ $\boxed{3}$-এ আসা যায় খালি $\boxed{2}$ থেকেই, এবং $\boxed{2}$-তে পৌঁছনোর পর 50% বার তুমি $\boxed{3}$-এ আসবে। সুতরাং প্যাটার্নটা হল এইরকম--কোনো একটা page নাও $\boxed{i}$. দ্যাখো কোন কোন page থেকে $\boxed{i}$-তে link আসছে। ধরো এদের নাম দিলাম $\boxed{j_1}$,...,$\boxed{j_k}$. Fig 8 দেখো। তাহলে equation-টা হবে--

$$p_i = (p_{j_1} + \cdots + p_{j_k}) \times \tfrac{1}{2}.$$

এখানে $\frac{1}{2}$ এসেছে, কারণ প্রত্যেকটা page থেকেই ঠিক 2-খানা করে link বেরিয়েছে। সুতরাং প্রত্যেকটা page-এর জন্যই আমরা একটা করে equation পাব। সব মিলিয়ে পাঁচটা equation হচ্ছে--

$$\begin{aligned}
p_1 &= \tfrac{1}{2} \times p_5 \\
p_2 &= \tfrac{1}{2} \times p_1 \\
p_3 &= \tfrac{1}{2} \times p_2 \\
p_4 &= \tfrac{1}{2}(p_1 + p_3) \\
p_5 &= \tfrac{1}{2}(p_2 + p_3 + p_4 + p_5).
\end{aligned}$$

Fig 8



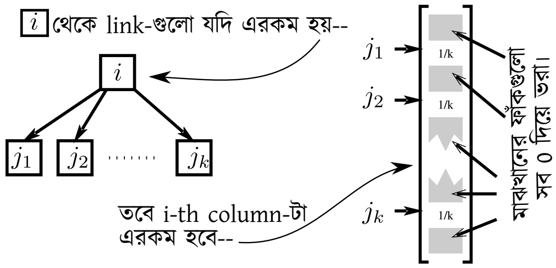


Fig 9

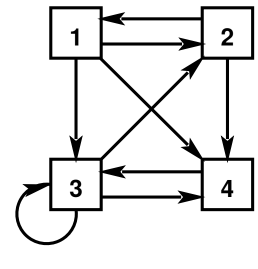


Fig 10

এই equation-গুলোকে একটু সাজিয়ে লিখলেই (ম্যাজিক!)--

$$\underbrace{\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix}}_{\vec{p}} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix}}_{\vec{p}}.$$

সুতরাং $\vec{p}$ হচ্ছে A-র একটা eigenvector, যেখানে eigenvalue-টা হল 1.
এই matrix-টাকে Fig 7 থেকে সরাসরি লিখে দেওয়া যায়। যতগুলো page আছে, তত সাইজের matrix হবে। যদি $\boxed{i}$ নম্বর page থেকে $\boxed{j_1}$,...,$\boxed{j_k}$ নম্বর page-গুলোতে link যায় (Fig 9), তবে i-th column-এর $j_1, ..., j_k$-th row-গুলোতে $\frac{1}{k}$ থাকবে, আর বাকিগুলোতে শূন্য থাকবে।

Example 12: Fig 10-এর জন্য matrix-টা কী হবে?

SOLUTION:

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 1 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{3} & 0 \end{bmatrix}$$

■

এই কায়দাটা বেশ general. তুমি যেকোনো-সংখ্যক page নিয়ে (মানে $\boxed{1}$,$\boxed{2}$, ইত্যাদি নিয়ে) তাদের মধ্যে যেভাবে খুশি link দাও, (মানে তীর চিহ্ন আঁকো) তবে সর্বদাই জানবে যে matrix-টার একটার eigenvalue হবে 1, এবং তার eigenvector-টা থেকে পাওয়া যাবে তুমি কোথায় কতক্ষণ সময় কাটাচ্ছ! অবশ্য এখানে আরো কিছু অংকের খুঁটিনাটি আছে, যেটা এখানে প্রাসঙ্গিক নয় বলে বাদ দিয়ে গেলাম, যেমন তীর চিহ্নগুলোর উপরে সামান্য কিছু শর্ত থাকলে সুবিধা হবে--তীর চিহ্নগুলো ধরে ধরে তুমি যেন যেকোনো জায়গা থেকে যেকোনো জায়গায় পৌঁছতে পারো, এবং অন্ততঃ একটা তীর চিহ্ন loop-এর মত হয় (মানে শুরু শেষ একই হয়, যেমন আমাদের উদাহরণে $\boxed{5}$ থেকে $\boxed{5}$-এ একটা loop ছিল)। এই দুটো শর্তই ইন্টারনেটের বেলায় পালিত হয়, এবং Google ঠিক এই অংকটাই করে। পুরো ইন্টারনেটের জন্য ওরা matrix-টা বানায় (পুরোটা বানাতে ওদের নাকি এক মাসের উপরে লেগে যায়), এর একটা eigenvalue যে 1 হবে সে তো জানাই, ওরা খালি বার করে তার eigenvector-টা। ব্যস্, তা থেকেই ওরা পেয়ে যায় কোন page কত গুরুত্বপূর্ণ। এই এক মাসের মধ্যে অবশ্য ইন্টারনেটে নতুন page এসে যায়, পুরনো link নষ্ট হয়ে যায়, তাই এই কাজটা ওদের ক্রমাগতই করে যেতে হয়।

Exercise 5: যদি Fig 7-এ একটা বাড়তি link থাকত $\boxed{2}$ থেকে $\boxed{1}$-এ, তবে matrix-টা কী হত? তার মানে এখন $\boxed{2}$ থেকে তিনটে তীরচিহ্ন বেরোচ্ছে। ■

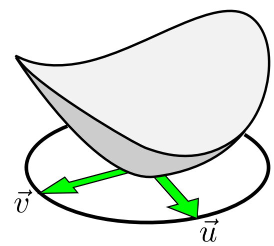


Fig 11

39.2 Quadratic form-এর সাথে সম্পর্ক

আমরা eigenvalue আর eigenvector-এর একটা প্রয়োগ দেখলাম। এটা বেশ একটু অদ্ভুত প্রয়োগ ছিল। Physics বা অন্যত্র eigenvector কিন্তু কাজে আসে সাধারণতঃ অন্য একটা কারণে। সেটা আসে quadratic form-এর সঙ্গে eigenvector আর eigenvalue-র একটা ওতপ্রোত সম্পর্ক থেকে। Quadratic form নিয়ে আমরা গত অধ্যায়েই বিস্তারিত আলোচনা করেছি। আজকের আলোচনাটুকু বোঝার জন্য অবশ্য অত কিছু লাগবে না। যেটুকু লাগবে সেটুকু মনে করিয়ে দিই।

Quadratic form-রা হল এক ধরণের real-valued function, যাদের domain হল $\mathbb{R}^n$. এদের চেহারা এইরকম--

$$f(\vec{x}) = \vec{x}' M \vec{x},$$

যেখানে M কোনো real, symmetric matrix.

যদি $n = 2$ নিই, তবে $f : \mathbb{R}^2 \to \mathbb{R}$, তাই এর গ্রাফ আঁকলে একটা surface-এর মত হবে। কম্পিউটারের সাহায্যে এঁকে Fig 11 পেয়েছি। ছবিটা দেখতে অনেকটা প্লেটের উপর কাপের মত হয়েছে। এখানে কাপটাই হল surface-টা, আর প্লেটের মত অংশটা আসলে xy-plane-এর উপরে unit disc-টা, মানে $\{(x, y) \ : \ x^2 + y^2 \leq 1\}$. লক্ষ্য কর যে, কাপটা কিন্তু সাধারণ কাপের মত সবদিকে সমানভাবে ওঠেনি। বরং কতকটা যেন নৌকার মত, একদিকে বেশী উঁচু, অন্য দিকে কম। সেটা বোঝানোর জন্য প্লেটটার গায় দুটো তীরচিহ্ন (মানে unit vector) এঁকেছি, $\vec{u}$ আর $\vec{v}$. এদের মধ্যে $\vec{u}$-তে surface-টা সবচেয়ে কম উঁচু , $\vec{v}$-তে সবচেয়ে বেশী, অর্থাৎ--

$$\forall \text{unit vector } \vec{x} \quad f(\vec{u}) \leq f(\vec{x}) \leq f(\vec{v}).$$

একটা quadratic form কোন্ unit vector-এর জন্য maximum minimum হয়, সেই প্রশ্নটা বিজ্ঞানের বিভিন্ন শাখাতেই কাজে লাগে। এর উত্তরে বেশ অপ্রত্যাশিতভাবেই eigenvalue আর eigenvector-র প্রসঙ্গ এসে পড়ে! বিস্তারিত উত্তরটা আমরা পরে শিখব। আপাততঃ মূল ব্যাপারটুকু বলি--

> যেহেতু M হল একটা real, symmetric matrix, তাই ওর সব eigenvalue-ই real হবে[4]। তাই ওদেরকে ছোটো থেকে বড় সাজানো যাবে। ধরো $\lambda_{\max}$ হল সবচেয়ে বড়টা, আর $\lambda_{\min}$ সবচেয়ে ছোটোটা। তবে যে unit vector-এর জন্য $f(\vec{x})$-টা maximum হবে সেটা $\lambda_{\max}$-এর জন্য একটা eigenvector হতে বাধ্য। এটা if and only if. যদি maximum চাও, তবে $\vec{x}$-কে $\lambda_{\max}$-এর জন্য eigenvector হতেই হবে, এবং বিপরীতপক্ষে এরকম যেকোনো eigenvector-এর জন্যই তুমি maximum পাবে। শুধু তাই নয়, এই maximum-টা হবে $\lambda_{\max}$-এর সমান! একইরকম ব্যাপার হবে minimum-এর জন্যও, খালি সেখানে $\lambda_{\min}$ নিয়ে কাজ করতে হবে। পুরো প্রমাণটা আসবে নবম অধ্যায়ে।

যেমন Fig 11-এর বেলায় আমরা বলতে পারি $\vec{v}$ হল M-এর সবচেয়ে বড় eigenvalue-র জন্য একটা eigenvector.

Exercise 6: ধরো M-এর দুটো eigenvalue-ই হল 2. ফলে দুটো eigenvalue-র মধ্যে কোনো ছোটোবড়োর ব্যাপার নেই। এবার বলো তো Fig 11-এর ছবিটা কীরকম দেখতে হবে! ■

---

[4]এটা এই অধ্যায়েই একটু পরে প্রমাণ করব।

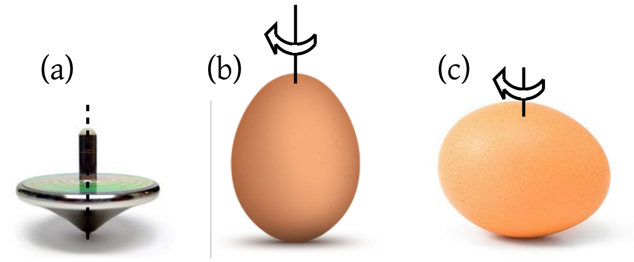


Fig 12

গত অধ্যায়েই বলেছি কী করে physics এবং statistics-এ quadratic form ব্যবহৃত হয়। আমাদের আজকের আলোচনা সেই একই সূত্র ধরে এগোবে। আমরা দেখব quadratic form-এর হাত ধরে কীভাবে physics এবং statistics-এ eigenvalue আর eigenvector-রা ঢুকে পড়ে।

39.2.1 Physics-এ প্রয়োগ

তুমি নিশ্চয়ই ছোটোবেলায় Fig 12(a)-এর মত লাট্টু ঘুরিয়েছ? লাট্টু নানারকম হয়, তবে সব লাট্টুরই মূল বৈশিষ্ট্য এই যে, একবার জোরে ঘুরিয়ে দিলে পরে ওরা খাড়া অবস্থাতেই ঘুরতে থাকে, পড়ে যায় না। ক্রমশঃ, ঘোরার বেগ যখন কমে আসে, তখন লাট্টুটা আর খাড়া থাকতে পারে না, কাত হয়ে পড়ে যায়। খাড়া হয়ে ঘোরাটা দেখতে ভালোই লাগে, কিন্তু ঘোরার সময়ে ওটা কিসের জোরে খাড়া থাকতে পারে, সেটা নিয়ে তলিয়ে দেখা আর হয় না। এই নিয়ে পণ্ডিতেরা অবশ্য বিস্তর গবেষণা করেছেন, এবং চমৎকার উত্তরও বার করেছেন। এবং সেই উত্তরের মধ্যে গুরুত্বপূর্ণ স্থান দখল করে আছে eigenvector-এর ধারণা! ব্যাপারটা সত্যিই মজার, একটু আলোচনা না করলেই নয়।
একটা লাট্টুকে ঘোরানোর সময়ে একটা বিশেষ axis বরাবরই ঘোরালেই ভালো কাজ হয়। কখনও অন্যভাবে ঘোরানোর চেষ্টা করে দেখেছ? তাতে মোটেই ঘোরে না। এই প্রসঙ্গে একটা ছোট্টো পরীক্ষা তুমি নিজেই করে দেখতে পারো। এর জন্য খালি একটা সিদ্ধ ডিম লাগবে।

> ডিমটাকে তুমি দুইভাবে লাট্টুর মত ঘোরানোর চেষ্টা করো, প্রথমে লম্বালম্বি Fig 12(b)-র মত, তারপর আড়াআড়ি Fig 12(c)-এর মত। দুই ক্ষেত্রেই আঙুলে সমান জোর দিলেও লক্ষ্য করবে দ্বিতীয় ক্ষেত্রে ডিমটা বেশীক্ষণ ঘুরছে। কেন?

কারণটা নানাভাবে বলা যায়, একটা সহজ কায়দা হল kinetic energy (গতিশক্তি) দিয়ে বলা। যখন তুমি লাট্টুটাকে ঘোরাচ্ছ, তখন তুমি ওর মধ্যে kinetic energy ঢুকিয়ে দিচ্ছ। যত বেশী kinetic energy ওটা নিতে পারবে, ততই ভালো করে বেশীক্ষণ ঘুরবে ওটা। ধরো তুমি ডিমটাকে (বা যেকোনো লাট্টুকে) ঘুরিয়ে দিলে। ঘুরিয়ে দেওয়ার angular velocity-কে $\vec{w}$ নাম দিলাম। মনে রেখো যে $\vec{w}$-এর মধ্যে দুটো জিনিস ঢুকে আছে--কত জোরে ঘুরিয়েছ, আর কোন axis বরাবর ঘুরিয়েছ। কত জোরে ঘোরাচ্ছ, সেটা তোমার আঙুলের জোরের উপর নির্ভর করে, আর axis-টা নির্ভর করে লাট্টুটাকে কীভাবে ধরেছ, তার উপরে। গত অধ্যায়েই আমরা দেখেছি যে, kinetic energy-টা হবে $\vec{w}$-এর একটা quadratic form,

$$\vec{w}'M\vec{w},$$

যেখানে M-কে আমরা লাট্টুটার moment of inertia matrix বলে থাকি। যত জোরে তুমি লাট্টুতে মোচড় দেবে, সেটা ততই বেশীক্ষণ ধরে ঘুরবে। এর মধ্যে নতুন কিছু নেই। কিন্তু তোমার আঙুলের জোর পরীক্ষা করা এখানে উদ্দেশ্য নয়। আমাদের গবেষণার বিষয় হল লাট্টুটা কোন্ দিক দিয়ে ঘোরালে সবচেয়ে ভাল ফল হবে, মানে ঘোরানোর axis-এর অভিমুখটা কী নিলে। তাই আমরা $\vec{w}$-কে একটা unit vector নেব, এবং চাইব যাতে এই quadratic form-টা maximum হয়। একটু আগেই বললাম যে, সেটা হবে যখন $\vec{w}$ হবে M-এর সবচেয়ে বড় eigenvalue-টার জন্য একটা eigenvector. এই axis-টা physics-এ এতই গুরুত্বপূর্ণ যে, এর একটা আলাদা নামই আছে--**principal axis of inertia.** আমরা যখন লাট্টু ঘোরাই, তখন না জেনেও আসলে এই principal axis of inertia বরাবরই ঘুরিয়ে থাকি।

39.2.2 Statistics-এ প্রয়োগ

Fig 13 দেখে বলো তো হাতে কী ধরা আছে! পেন? একটা কাঠি? ঠিক বোঝা যাচ্ছে না তো? ওই একই জিনিস একটু ঘুরিয়ে ধরা হয়েছে Fig 14-এ। এবার নিশ্চয়ই মোবাইল ফোনটাকে চিনতে কোনো কষ্ট হচ্ছে না? মোবাইলটা একটা 3D

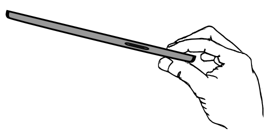


Fig 13

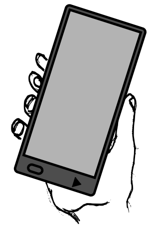


Fig 14

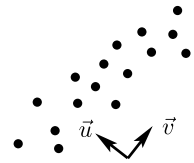


Fig 15

জিনিস, বইয়ের পাতাটা 2D. তাই যেভাবেই ধরি না কেন, ছবিতে খালি একটা পাশই ধরা পড়বে। কিন্তু কোন্ পাশটা দেখা যাচ্ছে, তার উপর নির্ভর করে 3D জিনিসটার প্রকৃত রূপটা কম বা বেশী বোঝা যাবে। অনেক সময়েই হয় যে, যে পাশটা বেশী ছড়ানো-ছেটানো, বিস্তৃত, সেই পাশটা বেশী তথ্যবহ হয়। মোবাইলের ছবি দুটোই তার উদাহরণ। কোন্ দিক থেকে ছবি নিলে বেশী ছড়ানো-ছেটানো লাগবে, সেটা বোঝা statistics-এর একটা গুরুত্বপূর্ণ কাজ, এবং সেই কাজে eigenvalue আর eigenvector-দের ব্যবহার হয়। সেটাই এবার সংক্ষেপে বলব।

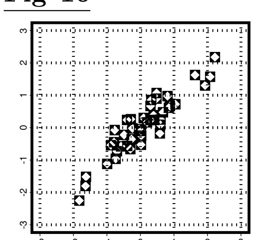
Fig 15-এ অনেকগুলো বিন্দু রয়েছে। আর দুটো দিক দেখানো আছে তীর চিহ্ন দিয়ে। যদি জিজ্ঞাসা করি এদের মধ্যে কোন্ দিক বরাবর বিন্দুগুলো সবচেয়ে বেশী ছড়িয়ে ছিটিয়ে আছে, তবে তোমার উত্তর নিশ্চয়ই হবে $\vec{v}$ বরাবর। এটা ছবি দেখে সহজেই বলা গেল। যদি বিন্দুগুলো খালি (x_i, y_i) হিসেবে দেওয়া থাকত, তবে কি স্রেফ অংক কষে তুমি এই দিকটা বার করতে পারতে? এইভাবে অংক কষে দিক বার করাটা অনেক সময়েই খুব কাজে দেয়, যেমন ধরো যদি বিন্দুগুলো 3-dimension-এ থাকত, তবে ছবি আঁকাটা সহজ হত না। যদি আরো বেশী dimension-এ থাকত, তবে তো ছবি আঁকা যেতই না, কিন্তু অংকের কায়দাটা তখনও কাজ করত। তাছাড়া চোখে দেখে করা গেলেও সেটা সবসময়ে সুবিধাজনক নয়। কারণ তার জন্য একটা মানুষকে দরকার, কিন্তু অংকের কায়দাটা একবার কম্পিউটারে প্রোগ্রাম করে নিলে বিভিন্নক্ষেত্রে বারবার প্রয়োগ করা যায়। এবার যা বলতে চলেছি, তাতে গত অধ্যায় থেকে একটা ধারণা লাগবে, variance-covariance matrix. চট্ করে একটু মনে করিয়ে দিই। Fig 15-এ যে বিন্দুগুলো রয়েছে, তারা সবাই আছে $\mathbb{R}^2$-তে। এই বিন্দুগুলো কীভাবে ছড়ানো-ছেটানো আছে, সেটা প্রকাশ করা যায় একটা 2×2 সাইজের matrix দিয়ে, ধরো V. যদি তুমি কোনো একটা দিক নাও (মানে একটা unit vector, ধরো $\vec{u}$), তবে সেই দিক বরাবর variance-টা হবে $\vec{u}'V\vec{u}$. যেহেতু এটা একটা quadratic form, তাই বুঝতেই পারছ যে, এটার maximum বার করতে গেলেই eigenvalue আর eigenvector এসে পড়বে। যদি আমরা $\vec{u}$-কে নিই V-এর সবচেয়ে বড় eigenvalue-এর জন্য কোনো eigenvector, অমনি সেই দিক বরাবর variance-টা maximum হবে। হাতেকলমে একটা অংক করলে ব্যাপারটা স্পষ্ট হবে।

Exercise 7: Fig 16-এ আবার অনেকগুলো বিন্দু আছে, একটা হেলানো দিক বরাবর ছড়িয়ে-ছিটিয়ে। এই দিকটা আমরা অংক কষে বার করতে চাই। এদের variance covariance matrix-টা আমি বলে দিচ্ছি--

$$\begin{bmatrix} 0.750 & 0.665 \\ 0.665 & 0.693 \end{bmatrix}.$$

এর দুটো eigenvalue বার কর। তার পর এদের মধ্যে যেটা বড়, তার জন্য একটা eigenvector বার কর। এবং সেটা একটা তীর চিহ্নের মত এঁকে দেখো হেলানো দিকটাই পাও কিনা।

Fig 16



HINT:
অংকটা যতটা জটিল মনে হচ্ছে, আসলে ততটা নয়। একটু ধরিয়ে দিচ্ছি। প্রথমে characteristic polynomial-টা বার কর। যেহেতু আমরা 2×2 নিয়ে কাজ করছি, তাই এটা একটা $\lambda^2+b\lambda+c$ জাতীয় জিনিস হবে। সুতরাং তার দুটো root অনায়াসেই বেরিয়ে যাবে (ক্যালকুলেটর লাগবে যদিও)। ধরো এদের মধ্যে বড়টাকে বললাম λ_1. যদি এর একটা eigenvector-কে $\vec{x}$ লিখি, তবে আমাদের সমাধান করতে হবে $(\lambda_1 I-A)\vec{x}=\vec{0}$. এর অনেক সমাধান সম্ভব, তার মধ্যে যেকোনো একটা $\vec{x}\neq\vec{0}$ নাও, এবং সেই দিক বরাবর unit vector-টা বার করে Fig 16-এর উপরে এঁকে দেখো।
হয়তো এত পরিশ্রম করতে তোমার ইচ্ছা করছে না। কিন্তু তাও অনুরোধ করব এ অংকটা বাদ না দিতে। Eigenvector দিয়ে যে সত্যিই কোনো কাজ হয়, তার এত প্রত্যক্ষ প্রমাণ অন্যভাবে পাওয়া কঠিন। ■

DAY 40
Algebraic এবং geometric multiplicity

আমরা দেখেছি যে, আমরা eigenvalue-র ধারণাটাকে দুভাবে ভাবা যায়। ধরো $A_{n\times n}$ একটা matrix. তবে কোনো $\lambda\in\mathbb{C}$-কে তার eigenvalue বলব যদি--

- এক, সেটা characteristic equation-এর একটা root হয়,
- আর দুই, যদি এমন $\vec{v}\neq\vec{0}$ পাওয়া যায়, যাতে $A\vec{v}=\lambda\vec{v}$ হয়।

এদের মধ্যে প্রথম কায়দাটা হল algebra-জাতীয় জিনিস, কারণ এখানে একটা equation solve করার ব্যাপার আছে। দ্বিতীয় কায়দাটা অনেকটা geometry-র মত, কারণ কোনো দিক বরাবর টেনে লম্বা করার ধারণা থেকে এই কায়দাটা জন্ম।
প্রথমে algebra-র কায়দাটার একটা উদাহরণ দেখি। ধরো একটা matrix আছে, যার characteristic polynomial হল $(\lambda-1)^3(\lambda-2)^4$. তার মানে ঠিক দুটোই eigenvalue আছে, 1 আর 2. লক্ষ্য করো, 1 এসেছে তিনবার (কারণ $(\lambda-1)$ রয়েছে $(\lambda-1)^3$ রূপে), আর 2 এসেছে চারবার (কারণ $(\lambda-2)$-এর মাথায় power হল 4)। এইভাবে characteristic polynomial-এ একটা eigenvalue কতবার এসেছে, তাকে বলে সেই eigenvalue-র **algebraic multiplicity**.

DEFINITION: **Algebraic multiplicity**

Let A be a square matrix over field $\mathbb{C}$. Let λ be an eigenvalue of A. Then its multiplicity as a root of the characteristic equation of A is called its **algebraic multiplicity** for A. We write it as $\mathrm{am}(\lambda, A)$.

একটা অংক কষে এই সংজ্ঞাটাকে হজম করে নিই--

Example 13: এই matrix-টার eigenvalue-গুলো বার করো। ওদের algebraic multiplicity-গুলোও বার করো।

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 3 & -1 \end{bmatrix}.$$

SOLUTION: এখানে A একটা lower triangular matrix, অতএব $\lambda I-A$-ও তাই। সুতরাং $\lambda I-A$-র diagonal entry-দের গুণ করে দিলেই determinant পাওয়া যাবে। মানে characteristic polynomial-টা হচ্ছে--

$$p(\lambda)=(\lambda-1)(\lambda-5)(\lambda-1)(\lambda+1)=(\lambda-1)^2(\lambda-5)(\lambda+1).$$

সুতরাং $p(\lambda) = 0$-র root-গুলো হল $1, 5$ আর -1. যেহেতু $(\lambda - 1)$-এর মাথায় power রয়েছে 2, তাই 1-এর algebraic multiplicity হবে 2. একইভাবে 5 আর -1 প্রত্যেকের algebraic multiplicity হবে 1. ■

এই গেল algebra-র দৃষ্টিভঙ্গী থেকে আলোচনা। এবার geometry-র দিক থেকে তাকানো যাক। ধরো $A_{n\times n}$-এর একটা eigenvalue হল λ. তার মানে এমন একটা $\vec{v} \neq \vec{0}$ আছে, যাতে $(\lambda I - A)\vec{v} = \vec{0}$ হয়। এরকম যাবতীয় $\vec{v}$-এর set-কে যে λ-র জন্য A-র eigenspace বলে, সেটা আগেই শুনেছ। অবশ্য eigenspace-টার মধ্যে $\vec{0}$-ও থাকে, যদিও সেটা eigenvector নয়। যাই হোক, eigenspace-টা হল $\mathbb{C}^n$-এর একটা subspace. সুতরাং তার একটা dimension আছে। এই dimension-টাকে বলে λ-র **geometric multiplicity**.
যেহেতু eigenspace-টা আসলে $\mathcal{N}(\lambda I - A)$, তাই geometric multiplicity-টা হল nullity$(\lambda I - A)$.

DEFINITION: **Geometric multiplicity**

Let A be a square matrix over field $\mathbb{C}$. Let $\lambda \in \mathbb{C}$ be an eigenvalue. Then the **geometric multiplicity** of λ for A is nullity$(\lambda I - A)$. We write it as gm(λ, A).

হাতে-কলমে একটা অংক করে সংজ্ঞাটা সড়গড় হয়ে নেওয়া যাক।

Example 14: Find the algebraic and geometric multiplicities of each eigenvalue of the matrix

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & -1 \\ 3 & 2 & -2 \end{bmatrix}$$

SOLUTION: প্রথমে characteristic polynomial-টা বার করি--

> Characteristic polynomial of A is
> $$|\lambda I - A| = \cdots = \lambda^3 - \lambda^2 - \lambda + 1 = (\lambda + 1)(\lambda - 1)^2.$$

দেখাই যাচ্ছে যে, characteristic equation-টার দুটো root, একজন -1, আসছে $(\lambda + 1)$ থেকে, অন্যজন 1, আসছে $(\lambda - 1)$ থেকে। যেহেতু $(\lambda - 1)$ রয়েছে $(\lambda - 1)^2$ হিসেবে, তাই 1-এর multiplicity হল 2. একই যুক্তিতে -1-এর multiplicity হল 1.

> The roots of this polynomial are -1 and 1 with multiplicities 1 and 2, respectively.
>
> $\therefore A$ has two eigenvalues $-1, 1$ with algebraic multiplicities $1, 2$, respectively.

এবার geometric multiplicity-গুলো বার করার পালা--

> Geometric multiplicity of -1 is nullity$(-I - A)$.

> Now
> $$-I-A=\begin{bmatrix} -1-1 & 1 & 0 \\ -1 & -1-2 & 1 \\ -3 & -2 & -1+2 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ -1 & -3 & 1 \\ -3 & -2 & 1 \end{bmatrix}.$$

এই matrix-টার nullity বার করতে হবে। নানাভাবে করা যায়--

- আমরা প্রথমে $(-I-A)\vec{x}=\vec{0}$ সমাধান করে $\mathscr{N}(-I-A)$-এর dimension বার করতে পারি।
- বা প্রথমে $(-I-A)$-এর rank বার করে nullity$(-I-A)=3-r(-I-A)$ করতে পারি, যেহেতু এখানে column-এর সংখ্যা 3.

এখানে আমরা দ্বিতীয় পথটা নেব, কারণ সেটা চোখে দেখেই করা যাবে--

> We observe that the sum of the first two rows is the third row. Also the second row is not a multiple of the first row. So $r(-I-A)=2$. Hence
> $$\text{nullity}(-I-A)=3-r(-I-A)=1.$$
> Thus geometric multiplicity of -1 is 1.

এই অংকটা শেষ হলেই আমরা আরেকটা কায়দা শিখব, যা দিয়ে আরো সহজে বলে দেওয়া যেত যে, এই geometric multiplicity-টা 1 হবে।

> Similarly, geometric multiplicity of 1 is nullity$(I-A)$. Now
> $$I-A=\begin{bmatrix} 1-1 & 1 & 0 \\ -1 & 1-2 & 1 \\ -3 & -2 & 1+2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & 1 \\ -3 & -2 & 3 \end{bmatrix}.$$

এটার rank বার করা খুবই সহজ। আমরা জানি যে এটার determinant হল 0 (কারণ 1 হল A-র একটা eigenvalue). সুতরাং rank অবশ্যই <3. প্রথম দুটো row কেউ কারো multiple নয়, সুতরাং rank অবশ্যই ≥ 2. অতএব $r(I-A)=2$ হতে বাধ্য।

> $\because 1$ is an eigenvalue, $\therefore |I-A|=0$, $\therefore r(I-A)<3$.
> Also the first two rows are not multiples of each other.
> $\therefore r(I-A)\geq 2$.
> $\therefore r(I-A)=2$.
> $\therefore$ nullity$(I-A)=3-r(I-A)=1$.
> So the geometric multiplicity of -1 is 1.

■

Exercise 8: Find the algebraic and geometric multiplicities of each eigenvalue of the matrix

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

■

আগের উদাহরণটায় আমরা দেখেছিলাম যে, -1-এর algebraic multiplicity ছিল 1. এটা থেকেই কিন্তু একবারে বলে দেওয়া যেত যে, geometric multiplicity-ও 1 হতে বাধ্য। তার কারণ একটা theorem আছে যে, সবসময়েই geometric multiplicity $\leq$ algebraic multiplicity হয়। আর geometric multiplicity কখনো 1-এর চেয়ে কম হতে পারে না। সুতরাং কোনো eigenvalue-র algebraic multiplicity যদি 1 হয়, অমনি বলে দেওয়া যায় যে, তার geometric multiplicity-ও 1 হবে।
Geometric multiplicity কেন ≥ 1 হয়, সেটা বোঝা কঠিন নয়। ধরো A-র একটা eigenvalue হল λ. তার মানে এমন একটা $\vec{v} \neq \vec{0}$ আছে, যাতে $(\lambda I - A)\vec{v} = \vec{0}$ হয়। সুতরাং $\mathscr{N}(\lambda I - A)$-র মধ্যে অন্ততঃ একটা non-null vector পেয়ে যাচ্ছি। তাই তার dimension তো আর শূন্য হতে পারে না, অন্ততঃ 1 হবেই!
এবার দেখি কেন $am(\lambda, A) \geq \text{gm}(\lambda, A)$ হতে বাধ্য।

> THEOREM
>
> For any eigenvalue of any matrix A,
>
> algebraic multiplicity $\geq$ geometric multiplicity.

Proof: প্রমাণটাতে বেশ একটু কৌশল আছে।

> Let α be an eigenvalue of A with geometric multiplicity g and algebraic multiplicity a.

দেখাব যে $a \geq g$.

> Let $\{\vec{v}_1, ..., \vec{v}_g\}$ be a basis of the eigenspace corresponding to eigenvalue α.

যেহেতু eigenspace-টার dimension হল g, তাই basis-এও ঠিক g-খানা vector আছে। এরা সকলেই A-র eigenvector, এবং সবার জন্যই eigenvalue হল সেই α.
এই vector-গুলো মিলে $\mathbb{C}^n$-এর মধ্যে একটা linearly independent set হয়। সুতরাং আমরা এটাকে পুরো $\mathbb{C}^n$-এর basis-এ extend করতে পারি (সেই যে extension theorem শিখেছিলাম দ্বিতীয় অধ্যায়ে, ইঁদুর বড় করতে করতে শুয়োর করা!)।

> We extend $\{\vec{v}_1, ..., \vec{v}_g\}$ to a basis $\{\vec{v}_1, ..., \vec{v}_g, \vec{v}_{g+1}, ..., \vec{v}_n\}$ of $\mathbb{C}^n$.

এখানে বাড়তি অংশটাকে $\vec{v}_{g+1}, ..., \vec{v}_n$ লিখেছি মানে এই নয় যে, আমরা $g < n$ ধরে নিচ্ছি। হতেই পারে যে $g = n$. সেক্ষেত্রে n-খানা vector পেয়েই গেছি, আর কোনো extension লাগবে না।
এবার এই n-খানা vector-কে পাশাপাশি column-এর মত দাঁড় করিয়ে একটা matrix বানাই।

> Let $P = [\ \vec{v}_1 \quad \cdots \quad \vec{v}_n\ ]$.
> Then P is a nonsingular matrix, since it is a square matrix with linearly independent columns.

এখানে $\vec{v}_1, ..., \vec{v}_g$-র সম্বন্ধে যে গুরুত্বপূর্ণ তথ্যটা আমাদের হাতে আছে, সেটা হল $A\vec{v}_i = \alpha\vec{v}_i$ হয় $(i = 1, ..., g)$. এই তথ্যটার সদ্ব্যবহার করতে হলে P-কে বাঁদিক থেকে A দিয়ে গুণ করা দরকার। এখানে প্রথম অধ্যায়[5] থেকে একটা ব্যাপার মনে করিয়ে দিই। যদি কোনো matrix দিয়ে কোনো matrix-কে বাঁদিক থেকে গুণ করো, তবে বাঁদিকের matrix-টা ডানদিকের matrix-টার প্রতিটা column-কে আলাদা করে গুণ করে। মানে, ডানদিকের matrix-টা যদি হয় $\left[\ \vec{b}_1 \quad \cdots \quad \vec{b}_p\ \right]$, তবে তাকে বাঁদিক থেকে কোনো L দিয়ে গুণ করলে হবে

$$L\left[\ \vec{b}_1 \quad \cdots \quad \vec{b}_p\ \right] = \left[\ L\vec{b}_1 \quad \cdots \quad L\vec{b}_p\ \right].$$

এইটা এবার কাজ লাগবে--

> Now
>
> $$\begin{aligned} AP &= [\ A\vec{v}_1 \quad \cdots \quad A\vec{v}_n\ ] \\ &= [\ \alpha\vec{v}_1 \quad \cdots \quad \alpha\vec{v}_g \quad A\vec{v}_{g+1} \quad \cdots \quad A\vec{v}_n\ ] \\ &= P\underbrace{\left[\begin{array}{cc} \alpha I_g & B \\ O & C \end{array}\right]}_{D}, \end{aligned}$$
>
> for some B, C.

মানে মোদ্দা কথা হল P-এর ডানদিকের matrix-টা দেখতে Fig 17-এর মত। ওই "যা খুশি" অংশটাকেই আমরা B আর C নাম ব্যবহার করে প্রকাশ করেছি। এই ধাপটাই সবচেয়ে কৌশলের। এই কৌশলটা আমরা বারবার এত ব্যবহার করব যে অভ্যাস হয়ে গেলে অতি সহজ লাগবে। কিন্তু প্রথমবার ভালো করে বুঝে নাও। যখন PD বার করছি তার column-গুলো হবে $P \times (D$-এর column-গুলো)। এবার D-এর প্রথম g-খানা column হল $\alpha\vec{e}_i$-এর মত। এখানে $\vec{e}_i$ বলতে সেই vector-টা বুঝছি যার সব entry-ই 0, খালি i-th entry-টা হল 1. গুণ করলেই দেখবে যে $P\vec{e}_i = \vec{v}_i$. এইটাই এখানে ব্যবহার করেছি।

> So $A = PDP^{-1}$.

এইবার আমরা দেখাব যে, A আর D-এর characteristic polynomial একই। এতে সুবিধা হবে, কারণ D হল block triangular.

[5] Block matrix operation আলোচনা করেছিলাম যেখানে।

Fig 17



The characteristic polynomial of A is

$$\begin{aligned} p_A(\lambda) &= |\lambda I - A| \\ &= |\lambda I - PDP^{-1}| \\ &= |P(\lambda I - D)P^{-1}| \end{aligned}$$

খুব ভালো করে বুঝে নাও কী করলাম। আমরা বাঁদিক থেকে P আর ডানদিক থেকে P^{-1} 'কমন' নিলাম। ফলে I-টা অপরিবর্তিত রয়েছে। চাইলে এটাকে দুইধাপে ভেঙে ভাবতে পারো--

$$\lambda I - PDP^{-1} = P(\lambda P^{-1} - DP^{-1}) = P(\lambda I - D)P^{-1}.$$

সুতরাং--

$$\begin{aligned} &= |P| \cdot |\lambda I - D| \cdot |P^{-1}| \\ &= |P| \cdot |\lambda I - D| \cdot |P^{-1}| \\ &= |\lambda I - D| \quad [\because |P| \cdot |P^{-1}| = 1] \\ &= \begin{vmatrix} (\lambda - \alpha)I_g & -B \\ O & \lambda I - C \end{vmatrix} \\ &= (\lambda - \alpha)^g p_C(\lambda), \end{aligned}$$

by law of determinant of block triangular matrices, where $p_C(\lambda) = |\lambda I - C|$ is the characteristic polynomial of C.

তার মানে $p_A(\lambda)$-র মধ্যে $(\lambda - \alpha)$ অন্তত g বার রয়েইছে ($p_C(\lambda)$-র মধ্যে আরো $(\lambda - \alpha)$ থাকতেও পারে, নাও পারে)। সুতরাং--

Since $(\lambda - p)^g$ is a factor of $p_A(\lambda)$, hence the multiplicity of α is $\geq g$. Thus $a \geq g$, as required.

[Q.E.D]

এই প্রসঙ্গে একটা নতুন শব্দ শিখে রাখি--

DEFINITION: **Regular eigenvalue**

An eigenvalue is called regular, if its algebraic multiplicity and geometric multiplicity are equal.

Exercise 9: ঠিক না ভুল বলো--

If $\text{am}(\lambda, A) = 1$ then λ must be a regular eigenvalue of A.

■

Exercise 10: এমন একটা diagonal matrix দিতে পারো, যার কোনো একটা eigenvalue-র algebraic multiplicity হল 3 আর সেই একই eigenvalue-এর geometric multiplicity হল 2?
■

DAY 41
বিভিন্ন ধর্ম (part 1)

41.1 কিছু বিশেষ ধরণের matrix

Eigenvalue বার করার একটা কায়দা আমরা দেখেছি--প্রথমে $|\lambda I - A|$ বার করতে হয়, যেটা সর্বদাই একটা polynomial হবে, তারপর দেখতে হবে সেটা কখন শূন্য হয়। এর মধ্যে প্রথম কাজটা বড় matrix-দের বেলায় যথেষ্ট বিরক্তিকর, আর দ্বিতীয় কাজটা তো প্রায় অসম্ভব রকমের কঠিন। সুতরাং কায়দাটা বাস্তবে নেহাত ছোটোখাটো চুনোপুঁটি matrix ছাড়া আর কোথাওই সহজে লাগানো যায় না। কিন্তু কিছু কিছু ভদ্রসভ্য matrix-এর ক্ষেত্রে কাজটা অনেকটাই সহজ। কারো কারো ক্ষেত্রে তো eigenvalue আর eigenvector-গুলো সবই অনায়াসে বের করে ফেলা যায়। আবার কোনো কোনো ক্ষেত্রে পুরো বার করা না গেলেও কিছু কিছু বৈশিষ্ট্য সহজেই জানা যায়। এবার এরকম কিছু ধরণের matrix-এর কথা শিখব।

41.1.1 Diagonal আর triangular

আমরা দেখেছি যে, diagonal বা triangular matrix-দের determinant বার করা সহজ। এদের diagonal-এ যে সংখ্যাগুলো আছে, তাদের গুণ করে দিলেই determinant পাওয়া যায়। সুতরাং এদের বেলায় eigenvalue আর eigenvector বার করাও সহজ।

Example 15: যদি $A = diag(1,2,3,4,4)$ হয়, তবে তার eigenvalue আর eigenvector-গুলো বার কর। সেই সঙ্গে ওদের eigenspace-গুলোও বার কর। ওদের algebraic multiplicity এবং geometric multiplicity-গুলো কত কত?

SOLUTION: $A = diag(1,2,3,4,4)$ মানে কী বুঝেছ তো? এর মানে

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}.$$

এখানে $\lambda I - A$ হবে একটা diagonal matrix–

$$\lambda I - A = \begin{bmatrix} \lambda-1 & 0 & 0 & 0 & 0 \\ 0 & \lambda-2 & 0 & 0 & 0 \\ 0 & 0 & \lambda-3 & 0 & 0 \\ 0 & 0 & 0 & \lambda-4 & 0 \\ 0 & 0 & 0 & 0 & \lambda-4 \end{bmatrix}.$$

সুতরাং $|\lambda I - A| = 0$ মানে হল

$$(\lambda-1)(\lambda-2)(\lambda-3)(\lambda-4)^2 = 0,$$

মানে eigenvalue-রা হল $1, 2, 3$ আর 4. প্রথম তিনজনের algebraic multiplicity হল 1, আর 4-এর algebraic multiplicity হচ্ছে 2. এবার eigenvector বার করার চেষ্টা করি। এর জন্য আমাদের সমাধান করতে হবে এই equation-টাকে--

$$(\lambda I - A)\vec{v} = \vec{0},$$

যেখানে λ হবে একটা eigenvalue.
যদি $\vec{v} = (x_1, ..., x_5)$ নিই, তবে

$$(\lambda I - A)\vec{v} = \begin{bmatrix} (\lambda-1)x_1 \\ (\lambda-2)x_2 \\ (\lambda-3)x_3 \\ (\lambda-4)x_4 \\ (\lambda-4)x_5 \end{bmatrix}.$$

এবার λ-র জায়গায় একে একে eigenvalue-গুলো বসানো যাক। যদি $\lambda = 1$ বসাই, তবে $(\lambda I - A)\vec{v} = \vec{0}$ মানে $x_2 = \cdots = x_5 = 0$. সুতরাং খালি x_1-টাই $\neq 0$ হতে পারে। অতএব $\lambda = 1$-এর জন্য A-র একটা eigenvector হবে $\vec{e_1} = (1, 0, 0, 0, 0)$. একই যুক্তিতে $\lambda = 2$-এর জন্য একটা eigenvector হবে $(0, 1, 0, 0, 0)$, আর $\lambda = 3$-এর জন্য হবে $(0, 0, 1, 0, 0)$. যদি $\lambda = 4$ নিই, তবে ব্যাপারটা সামান্য অন্যরকম। যেহেতু 4 দুইবার এসেছে, তাই $\vec{e_4} = (0, 0, 0, 1, 0)$ এবং $\vec{e_5} = (0, 0, 0, 0, 1)$ দুজনেই এক্ষেত্রে eigenvector হবে।
Eigenspace-গুলোও বেরিয়ে যাচ্ছে সহজেই, $\lambda = 1$-এর জন্য $\text{span}\{\vec{e_1}\}$, আর $\lambda = 2$-এর জন্য $\text{span}\{\vec{e_2}\}$, আর $\lambda = 3$-এর জন্য $\text{span}\{\vec{e_3}\}$. যখন $\lambda = 4$ নেব, তখন eigenspace-টা হবে $\text{span}\{\vec{e_4}, \vec{e_5}\}$.
Geometric multiplicity-গুলো হচ্ছে, যথাক্রমে, 1, 1, 1 এবং 2. ■

এই অংকটা যদি বুঝে থাকো, তবে নীচের theorem-টা অপ্রত্যাশিত লাগা উচিত নয়--

THEOREM

If $A = diag(a_1, ..., a_n)$, then its eigenvalues are $a_1, ..., a_n$. Also $\vec{e_i}$ is an eigenvector corresponding to the eigenvalue a_i. The algebraic multiplicity and geometric multiplicity of an eigenvalue are both same and equal to the number of times the eigenvalue is repeated in the diagonal.

Exercise 11: একটা diagonal matrix দিলাম $A = \text{diag}(a_1, ..., a_9)$. যদি $\vec{e_1} + \vec{e_2}$ এর একটা eigenvector হয়, তবে দেখাও যে $a_1 = a_2$ হতে বাধ্য। ■

Exercise 12: ধরো $A = \text{diag}(a_1, ..., a_n)$. বলে দেওয়া আছে যে $\vec{e_1} + \vec{e_2}$ হল এর একটা eigenvector, যার জন্য eigenvalue-টা হল λ_1. একইভাবে এও বলা আছে যে $2\vec{e_2} + 3\vec{e_3}$ আরেকটা eigenvector, যার জন্য eigenvalue হল λ_2. তাহলে $\lambda_1 - \lambda_2$ বার কর।
HINT: আগের অংকের যুক্তিটা এখানেও কাজ লাগবে। ■

আমরা জানি diagonal-দের গুণ করে determinant বার করার কৌশলটা শুধু diagonal matrix-দের একচেটিয়া সম্পত্তি নয়, ওটা যেকোনো triangular matrix-এর বেলাতেই খাটে। সুতরাং এক্ষেত্রেও diagonal-এর সংখ্যাগুলোই যে eigenvalue হবে সন্দেহ নেই। দুঃখের কথা eigenvector-এর বেলায় এক্ষেত্রে তেমন কিছুই বলা যায় না। কোন eigenvalue-টা diagonal-এ কতবার এসেছে, সেটাই তার algebraic multiplicity. কিন্তু geometric multiplicity নিয়ে তেমন কিছুই বলা যায় না!

41.1.2 Nonsingular

Nonsingular matrix-দের বেলায় eigenvalue বার করার কোনো এককথায় সহজ কায়দা জানা নেই। কিন্তু একটা জিনিস বলা যায়--eigenvalue-রা সবাই nonzero হতে বাধ্য। এর উল্টো দিকটাও সত্যি-- একটা matrix-এর সবগুলো eigenvalue যদি nonzero হয়, তবে matrix-টা nonsingular হতে বাধ্য। এটাই নীচের theorem-টার বক্তব্য--

> **THEOREM**
>
> A square matrix is nonsingular if and only if all its eigenvalues are nonzero.

Proof: প্রমাণটা খুবই সহজ। λ একটা eigenvalue হওয়া মানে $|\lambda I - A| \neq 0$. সুতরাং $\lambda = 0$ হওয়া মানে দাঁড়াচ্ছে $|-A| = 0$, অর্থাৎ $|A| = 0$ (কারণ মনে আছে নিশ্চয়ই যে $|-A| = (-1)^n|A|$, যেখানে A হল $n \times n$). অংকের ভাষায় লিখে ফেললে--

> $A_{n\times n}$ has all eigenvalues nonzero
>
> $\iff |0I - A| = |-A| = (-1)^n|A| \neq 0 \iff |A| \neq 0 \iff A$ is nonsingular.

[Q.E.D]

Exercise 13: ধরো A কোনো একটা singular matrix, যার rank হল r, এবং যাবতীয় nonzero eigenvalue-র মোট সংখ্যা হল k. এখানে eigenvalue গোণার সময়ে algebraic multiplicity ধরে গুণতে হবে, যেমন characteristic polynomial যদি $(\lambda - 2)^4\lambda^3$ হয়, তবে nonzero eigenvalue-র সংখ্যা হবে 4. এবার নীচের বাক্য কয়টা লক্ষ্য কর। এদের মধ্যে কোন্‌টা যেকোনো square matrix-এর ক্ষেত্রেই খাটে?

(A) $k \leq r$ (B) $k = r$ (C) $k \geq r$ (D) কোনোটাই নয়।

■

41.1.3 Unitary আর (real) orthogonal

যে সব (complex) matrix A-র জন্য $A^*A = AA^* = I$ হয়, তাদের যে unitary বলে, সেটা আশা করি ভুলে যাও নি। যদি A-টা তার উপরে আবার একটা real matrix-ও হয়, তবে $A^* = A'$. এক্ষেত্রে তাই শর্তটা হবে $A'A = AA' = I$. তখন আমরা এদের বলি orthogonal matrix, কখনো বা real কথাটা ভালো করে মনে করিয়ে দেবার জন্য বলি real orthogonal matrix.

এদের eigenvalue বিষয়ে একটা কথা বলা যায়, eigenvalue-গুলোর modulus অবশ্যই 1 হতে বাধ্য। যেমন $\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i$ হতে, পারে, বা i হতে পারে, বা 1-ও হতে পারে। কিন্তু কোনোভাবেই $1+i$ হতে পারে না, কারণ তার modulus হল $\sqrt{1^2+1^2} = \sqrt{2} \neq 1$.

> **THEOREM**
>
> The eigenvalue of a unitary matrix (and hence of a real orthogonal matrix) must have modulus 1.

প্রমাণটা খুবই ছোটো, কিন্তু একটা কৌশল আছে। সেটা হল এই যে, যদি তুমি কোনো $\vec{v} \in \mathbb{C}^n$ নাও, তবে $\vec{v}^*\vec{v}$ সর্বদা একটা nonnegative real number হবে। কারণ, ধরো $\vec{v}$-এর entry-রা হল $z_1, ..., z_n \in \mathbb{C}$. তবে

$$\vec{v}^*\vec{v} = \begin{bmatrix} \overline{z}_1 & \cdots & \overline{z}_n \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \sum |z_i|^2 \geq 0,$$

কারণ যেকোনো complex সংখ্যারই modulus হয় একটা nonnegative real number. আমরা আরো জানি যে, $|z_i|^2 = 0$ হওয়া আর $z_i = 0$ হওয়া সমার্থক। তাই আমরা বলতে পারি যে, $\vec{v}^*\vec{v} = 0$ হওয়া আর $\vec{v} = \vec{0}$ হওয়া সমার্থক (ভেবে দেখো!)।

Theorem-এর প্রমাণটা শুরু করার আগে আরেকটা জিনিস মনে করিয়ে দিই। যদি $\alpha \in \mathbb{C}$ হয়, তবে যেকোনো matrix-এর জন্যই $(\alpha A)^* = \overline{\alpha} A^*$ হবে।

এবার তবে প্রমাণটায় হাত দেওয়া যাক।

Proof:

> Let A be any unitary matrix. Let $\lambda \in \mathbb{C}$ be any eigenvalue with a corresponding eigenvector $\vec{v} \neq \vec{0}$.

এবার আমরা $(A\vec{v})^*(A\vec{v})$-কে দুইভাবে লিখব, প্রথমবার $A\vec{v} = \lambda\vec{v}$ ব্যবহার করে, দ্বিতীয়বার সরাসরি--

> Then
>
> $$\begin{aligned} (A\vec{v})^*(A\vec{v}) &= (\lambda\vec{v})^*(\lambda\vec{v}) \\ &= \overline{\lambda}\lambda(\vec{v}^*\vec{v}) \\ &= |\lambda|^2(\vec{v}^*\vec{v}). \end{aligned}$$

এবার সরাসরি এগোব। মনে রেখো যেকোনো M, N-এর জন্যই $(MN)^* = N^*M^*$ হয়।

> Again,
>
> $$\begin{aligned} (A\vec{v})^*(A\vec{v}) &= \vec{v}^*A^*A\vec{v} \\ &= \vec{v}^*I\vec{v} \quad [\because A \text{ is unitary}] \\ &= \vec{v}^*\vec{v}. \end{aligned}$$

আমরা $(A\vec{v})^*(A\vec{v})$-কে দুইভাবে দলাইমলাই করে দুটো জিনিস পেয়েছি। তারা একই জিনিসের দুটো রূপ, তাই সমান হতে বাধ্য--

> Hence
>
> $$|\lambda|^2\vec{v}^*\vec{v} = \vec{v}^*\vec{v}.$$
>
> $\because \vec{v} \neq \vec{0} \therefore \vec{v}^*\vec{v} \neq 0.$
>
> So $|\lambda| = 1,$ as required $(\because |\lambda| \geq 0)$.

[Q.E.D]

41.1.4 Hermitian, symmetric, skew Hermitian, skew symmetric

এবার আমরা আরেকধরণের matrix দেখব, যাদের বেলায় eigenvalue বার করা সহজ না হলেও, অন্ততঃ এটুকু সহজেই বলা যায় যে, eigenvalue-রা সবাই real সংখ্যা হবে। এরা হল Hermitian matrix, মানে যাদের জন্য $A^* = A$ হয়। যদি A-র সব entry-ই real হয়, তবে সেই বিশেষ ক্ষেত্রে A-টা হবে real, symmetric matrix. সবগুলো eigenvalue-ই real হওয়া খুবই বড় খবর, কারণ আমরা আগেই দেখেছিলাম যে, eigenvalue-রা real হলেই খালি ওদেরকে "টেনে লম্বা করা"-র উপমা দিয়ে ভাবা যায়। যদি real না হয়, তবে ওদেরকে নিয়ে চিন্তা করা অপেক্ষাকৃত কঠিন।

> **THEOREM**
>
> All the eigenvalues of a Hermitian matrix are real.

খালি Real কেসটার প্রমাণ লাগবে নীচের অংকে।

Example 16: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). For a real, symmetric matrix, eigenvalues are

(A) all real

(B) all complex

(C) cannot be ascertained

(D) both real and complex

(2021.CC6.1k)

SOLUTION: এই অংকটায় ভাষার সমস্যা আছে। মনে রেখো যে, complex সংখ্যাদের মধ্যে real সংখ্যাদেরও ধরা হয়। তাই "3 একটা real সংখ্যা" এই বাক্যটাও যেমন সঠিক, তেমনি "3 একটা complex সংখ্যা" এই বাক্যটাও ততটাই সঠিক। সেই যুক্তিতে এখানে (A), (B) এবং (D) তিনটে option-ই সঠিক। কিন্তু প্রশ্নের ভাষা থেকে বোঝা যাচ্ছে যে, এখানে একটাই option সঠিক হওয়ার কথা। তার মানে প্রশ্নকর্তা এখানে উত্তর প্রত্যাশা করছেন খালি (A). কিন্তু অংকের ব্যাকরণ অনুযায়ী (B) আর (D)-ও সমান দাবীদার।

> Let λ be an eigenvalue of a real symmetric matrix, A.
>
> Then $\exists \vec{v} \neq \vec{0} \quad A\vec{v} = \lambda\vec{v}$
>
> Multiplying from left by $\vec{v}^*$,
>
> $$\vec{v}^* A\vec{v} = \lambda\vec{v}^*\vec{v}. \qquad (*)$$
>
> Taking conjugate transpose of both sides,
>
> $$\vec{v}^* A\vec{v} = \bar{\lambda}\vec{v}^*\vec{v}, \qquad (**)$$
>
> Since $A^* = A$.

এখানে অবাক লাগতে পারে যে, আমরা তো খালি real কেসটার প্রমাণ করছি, তাহলে conjugate transpose নিচ্ছি কেন? তার কারণ আমরা এখনও দেখাই নি যে λ-টা একটা real number, আর তাছাড়া eigenvector-টার entry-গুলো real নাই হতে পারে।

> So, from (*) and (**), we get
>
> $$\lambda\vec{v}^*\vec{v} = \bar{\lambda}\vec{v}^*\vec{v}.$$
>
> Now, since $\vec{v} \neq \vec{0}$, hence $\lambda = \bar{\lambda}$.

এর মানেই হল $\lambda \in \mathbb{R}$, কারণ $\lambda = x + iy \quad (x, y \in \mathbb{R})$ হলে $\bar{\lambda} = x - iy$. সুতরাং $\lambda = \bar{\lambda}$ হওয়া মানে $y = -y$ হওয়া, অর্থাৎ $y = 0$.

> So λ must be real, as required.

■

Theorem-টার real অংশটা প্রমাণ হল। নীচের অংকে পুরো theorem-টারই প্রমাণ চেয়েছে।

Exercise 14: Show that the eigenvalues of a Hermitian matrix are all real.[2] **(2022.CC6.3fii)**
■

একইরকম কাজ করা যায় real skew symmetric matrix-দের বেলাতেও। আশা করি মনে আছে যে, skew symmetric মানে এমন A, যার জন্য $A' = -A$ হয়। একইভাবে কোনো complex matrix-কে **skew Hermitian** বলা হবে যদি $A^* = -A$ হয়। একটা real skew symmetric matrix অবশ্যই একটা complex skew Hermitian matrix-ও হবে। এদের নিয়েই পরের অংকটা।

Exercise 15: Prove that the eigenvalues of a skew Hermitian matrix are purely imaginary or zero.
HINT:
আগের অংকের অনুকরণে এগোও।
■

Example 17: If λ is an eigenvalue of a real skew symmetric matrix, prove that

$$\left|\frac{1-\lambda}{1+\lambda}\right| = 1.$$

SOLUTION: এই অংকটা ছবি দিয়ে ভাবার পক্ষে ভালো। আমরা আগের অংকেই দেখেছি যে, skew symmetric matrix-দের eigenvalue-রা purely imaginary হয়। অতএব λ-কে মনে করতে পারো complex plane-এর imaginary axis-এর উপরে একটা বিন্দু। আমাদের যেটা দেখাতে বলেছে, সেটাকে এইভাবে লিখে নেওয়া যায়--

$$|\lambda - 1| = |\lambda - (-1)|.$$

আশা করি মনে আছে যে $w, z \in \mathbb{C}$ হলে $|w - z|$ মানে complex plane-এ w আর z বিন্দুদুটোর মধ্যেকার দূরত্ব। সুতরাং Fig 18-এর দুটো দূরত্বকে সমান দেখানোটাই আমাদের কাজ। দূরত্ব দুটো যে সমান সেটা বুঝতেই পারছ, কারণ 1 এবং -1 রয়েছে imaginary axis-এর দুই দিকে সমান দূরত্বে। বস্তুতঃ $1, -1$-র বদলে যেকোনো $a \in \mathbb{R}$-এর জন্য a আর $-a$ নিয়েও অংকটা হতে পারত।

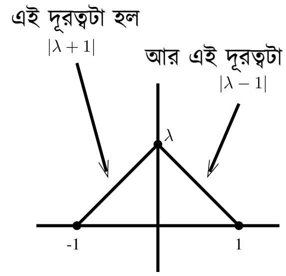


Fig 18

$\because A$ is skew symmetric, $\therefore \lambda$ must be purely imaginary.
Let $\lambda = iy$ for some $y \in \mathbb{R}$.
Then $|1+\lambda| = |1+iy| = \sqrt{1+y^2}$.
Again $|1-\lambda| = |1-iy| = \sqrt{1+y^2}$.
So $|1+\lambda| = |1-\lambda|$.
Also $(1+y^2) \neq 0$. So dividing by $|1+\lambda|$ we have

$$\left|\frac{1-\lambda}{1+\lambda}\right| = \frac{|1-\lambda|}{|1+\lambda|} = 1,$$

as required.

■

DAY 42
বিভিন্ন ধর্ম (part 2)

42.1 Determinant, trace-এর সাথে সম্পর্ক

Determinant আর trace-এর সঙ্গে eigenvalue-দের একটা সম্পর্ক আছে, যেটা বেশ অপ্রত্যাশিত। সবগুলো eigenvalue-কে গুণ করলে determinant পাবে, আর যোগ করলে trace! একটা উদাহরণ দেখে নেওয়া যাক।

Example 18: এই matrix-টার বেলায় পরীক্ষা করে দেখো--

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

.

SOLUTION: এখানে characteristic polynomial হল

$$|\lambda I - A| = (\lambda-1)(\lambda-4) - 6 = \lambda^2 - 5\lambda - 2.$$

তাই eigenvalue দুটো হবে $\frac{1}{2}(5 \pm \sqrt{33})$. এদেরকে যোগ করলে পাচ্ছি 5. এবং matrix-টার trace-ও ঠিক সেটাই, $1+4=5$. গুণ করলে হয়

$$\tfrac{1}{4}(5^2 - 33) = -2,$$

এবং matrix-টার determinant-ও ঠিক সেটাই। ■

এখানে একটা কথা বলে রাখি--যখন গুণ বা যোগ করছ, তখন কিন্তু একটা eigenvalue যতবার আসে, ততবার হিসেবে নিতে হবে, মানে algebraic multiplicity যত, ততবার। একটা উদাহরণ দেখাচ্ছি।

Example 19: মনে করো একটা matrix-এর characteristic polynomial হল $(\lambda-2)^3(\lambda+3)$. তাহলে matrix-টার determinant আর trace বার কর।

SOLUTION: এখানে eigenvalue-রা হল $2,2,2$ আর -3. লক্ষ্য কর 2-টা তিনবার লিখেছি, কারণ ওর algebraic multiplicity হল 3. সুতরাং determinant হবে

$$2\times 2\times 2\times(-3)=-24,$$

আর trace হবে

$$2+2+2+(-3)=3.$$

■

কেন এমনটা হয়, সেটা বোঝার জন্য একটা জিনিস মাথায় রাখতে হবে। যদি একটা polynomial দিই

$$x^n+c_{n-1}x^{n-1}+\cdots+c_1x+c_0,$$

যার root-গুলো হল $r_1,...,r_n\in\mathbb{C}$, তবে

$$\begin{aligned} c_0 &= (-1)^n r_1r_2\cdots r_n,\\ c_{n-1} &= -(r_1+r_2+\cdots+r_n). \end{aligned}$$

যেমন এই polynomial-টার জন্য (যেখানে $n=2$)--

$$x^2\underbrace{-5}_{c_1}x+\underbrace{6}_{c_0}=(x-\underbrace{2}_{r_1})(x-\underbrace{3}_{r_2})$$

$c_0=6=(-1)^2\times 2\times 3,$ আর $c_{n-1}=c_1=-5=-(2+3)$.
এই জিনিসটাই আমরা লাগাব characteristic polynomial-এর বেলায়। যদি এই polynomial-টাকে লিখি

$$|\lambda I-A|=\lambda^n+c_{n-1}\lambda^{n-1}+\cdots+c_1\lambda+c_0,$$

তবে c_0 কী হবে? খুব সহজ উত্তর, $\lambda=0$ বসিয়ে দিলেই পাবে $c_0=|-A|=(-1)^n|A|$. এদিকে characteristic polynomial-এর root-গুলোই তো eigenvalue-রা। সুতরাং $(-1)^n|A|=(-1)^n\lambda_1\lambda_2\cdots\lambda_n$. এবার দুদিক থেকে $(-1)^n$ কেটে নিলেই হল।
Trace-এর বেলাতেও একই রকম যুক্তি--এখানে a_{n-1} বার করতে হবে, মানে $|\lambda I-A|$-র মধ্যে λ^{n-1}-এর coefficient. একটা উদাহরণ নিলে সুবিধা হবে, ধরো $n=4$. তার মানে

$$|\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & -a_{13} & -a_{14}\\ -a_{21} & \lambda-a_{22} & -a_{23} & -a_{24}\\ -a_{31} & -a_{32} & \lambda-a_{33} & -a_{34}\\ -a_{41} & -a_{42} & -a_{43} & \lambda-a_{44}\end{vmatrix}.$$

আমাদের বার করতে হবে λ^3-এর coefficient. লক্ষ্য কর যে, λ-ওয়ালা জিনিসগুলো সবাই আছে diagonal-এ। সুতরাং determinant-টার মধ্যে যে 4!-খানা term থাকবে, তাদের মধ্যে λ^3 থাকবে খালি সেগুলোতেই যেগুলোতে তিনখানা entry

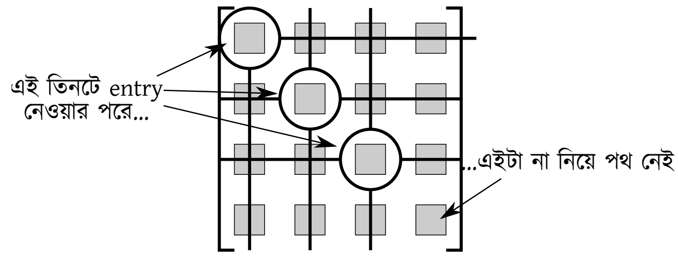


Fig 19

এসেছে diagonal থেকে। এদিকে determinant-এর প্রতিটা term-এ প্রত্যেকটা row আর column থেকেই ঠিক একটা করেই entry থাকে, সুতরাং চতুর্থ entry-টাকেও বাধ্য হয়েই diagonal থেকেই আসতে হবে (Fig 19)। তার মানে এই term-টা হবে

$$(\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}),$$

যার মধ্যে λ^3-এর coefficient হবে $-(a_{11} + a_{22} + a_{33} + a_{44})$. ঠিক একই যুক্তিতে $n \times n$ সাইজের matrix-এর বেলায় λ^{n-1}-এর coefficient হবে $c_{n-1} = -\sum a_{ii} = -tr(A)$. আবার আমরা জানি যে $c_{n-1} = -\sum \lambda_i$. সুতরাং পেয়ে যাচ্ছি $tr(A) = \sum \lambda_i$, ঠিক যেমনটা দরকার ছিল।

Exercise 16: একটা 3×3 সাইজের matrix-এর trace আর determinant দুইই হল 6. বলা আছে যে, একটা eigenvalue হল 1. বাকি eigenvalue দুটো কী কী? ■

42.2 Linear independence

এতক্ষণ eigenvalue-দের কথাই বেশী হল, এবার eigenvector-দের দিকে একটু চোখ ফেরানো যাক। ধরো A একটা square matrix, যার একটা eigenvalue হল λ, আর $\vec{v}$ হল তার জন্য একটা eigenvector. তবে এতে নিশ্চয়ই সন্দেহ নেই যে, $2\vec{v}$-ও হবে ওই λ-র জন্যই আরেকটা eigenvector, কারণ ওই 2-টা "বেরিয়ে আসবে"--

$$A(2\vec{v}) = 2A\vec{v} = 2\lambda\vec{v} = \lambda(2\vec{v}).$$

একই যুক্তি খাটত 2-এর বদলে যেকোনো nonzero সংখ্যা নিলে। সুতরাং যদি তোমাকে দুটো eigenvector দিই, যাদের জন্য eigenvalue দুটো আলাদা, তবে সেইরকম দুটো eigenvector পরস্পরের multiple হতে পারে না, মানে তারা linearly independent হতে বাধ্য। এই জিনিসটাই নীচের অংকে কাজে লাগবে, যদিও লিখব একটু অন্য ঢঙে।

Example 20: Show that two eigenvectors of a square matrix A over a field F corresponding to two distinct eigenvalues of A are linearly independent.[2] **(2019.CC6.3cii)**

SOLUTION:

> Let $\mu \neq \lambda$ be two eigenvalues of a matrix A.
>
> Let $\vec{u}$ and $\vec{v}$ be two corresponding eigenvectors.
>
> Let
>
> $$a\vec{u} + b\vec{v} = \vec{0}. \qquad (*)$$
>
> Shall show that $a = b = 0$.

Multiplying by A from left, $A(a\vec{u}+b\vec{v})=\vec{0}$, or $aA\vec{u}+bA\vec{v}=\vec{0}$, ie,

$$a\mu\vec{u}+b\lambda\vec{v}=\vec{0}. \quad (**)$$

Subtracting μ times (*) from (**),

$$b(\lambda-\mu)\vec{v}=\vec{0}.$$

Hence $b=0$, since $\vec{v}\neq\vec{0}$ and $\lambda\neq\mu$.
So, from (*), we get $a\vec{u}=\vec{0}$, and hence $a=0$, since $\vec{u}\neq\vec{0}$.

■

একইরকম কথা খাটে তিনটে, বা তার চেয়েও বেশী সংখ্যক eigenvector-এর বেলাতেও--

THEOREM

Let $\vec{v}_1,...,\vec{v}_k$ be eigenvectors of a matrix A corresponding to *distinct* eigenvalues $\lambda_1,...,\lambda_k$. Then $\{\vec{v}_1,...,\vec{v}_k\}$ must be a linearly independent set.

Proof:

We shall use induction on k.

Induction দিয়ে প্রমাণের তিনটে অংশ হয়--basis, hypothesis আর step. প্রথমে basis, সেখানে আমরা $k=1$-এর জন্য সরাসরি দেখাব।

Basis: If $k=1$ then $\{\vec{v}_1\}$ is a linearly independent set, since $\vec{v}_1\neq\vec{0}$.

এরপরে hypothesis, যেখানে আমরা $k=1,...,m$-এর জন্য theorem-টা প্রমাণ হয়ে গেছে ধরে নেব।

Hypothesis: Assume the result for $k=1,...,m$.

আসল কাজটা হবে তৃতীয় অংশে। এখানে আমরা $k=m+1$-এর জন্য প্রমাণটা করব।

Step: Shall show for $k=m+1$.

Induction hypothesis থেকে দেওয়া আছে যে $\{\vec{v}_1,...,\vec{v}_m\}$ হল linearly independent. দেখাতে হবে যে $\{\vec{v}_1,...,\vec{v}_{m+1}\}$-ও linearly independent. তার মানে $c_1\vec{v}_1+\cdots+c_{m+1}\vec{v}_{m+1}=\vec{0}$ হলে $c_1=\cdots=c_{m+1}=0$ দেখাব।

Let

$$c_1\vec{v}_1 + \cdots + c_{m+1}\vec{v}_{m+1} = \vec{0} \qquad (*)$$

for some $c_1, ..., c_{m+1} \in \mathbb{C}$.
Shall show $c_1 = \cdots = c_{m+1} = 0$.
Multiplying by A from left,

$$c_1 A\vec{v}_1 + \cdots + c_{m+1} A\vec{v}_{m+1} = \vec{0}$$

or

$$\lambda_1 c_1 \vec{v}_1 + \cdots + \lambda_{m+1} c_{m+1} \vec{v}_{m+1} = \vec{0}. \qquad (**)$$

এটা হল কারণ $A\vec{v}_i = \lambda_i \vec{v}_i$.
খালি যদি $\{\vec{v}_1, ..., \vec{v}_m\}$ থাকত, তবে তো induction hypothesis থেকেই কাজ হয়ে যেত। সমস্যা হয়েছে শেষের ওই $\lambda_{m+1}\vec{v}_{m+1}$-টা এসে গিয়েই। ওটাকে কায়দা করে তাড়াতে হবে। কায়দাটা আমরা আগের অংকের সমাধানেই দেখেছি।

Multiplying $(*)$ by λ_{m+1} we have

$$c_1\lambda_{m+1}\vec{v}_1 + \cdots + c_{m+1}\lambda_{m+1}\vec{v}_{m+1} = \vec{0}.$$

Subtracting from $(**)$,

$$c_1(\lambda_1 - \lambda_{m+1})\vec{v}_1 + \cdots + c_m(\lambda_m - \lambda_{m+1})\vec{v}_m = \vec{0}.$$

By induction hypothesis $\{\vec{v}_1, ..., \vec{v}_m\}$ is linearly independent.

$$c_1(\lambda_1 - \lambda_{m+1}) = \cdots = c_m(\lambda_m - \lambda_{m+1}) = 0.$$

এখান থেকে $c_1 = \cdots = c_m = 0$ দেখাব--

Since λ_i's are all distinct,

$$\therefore \forall i = 1, ..., m \quad (\lambda_i - \lambda_{m+1}) \neq 0.$$

So $\therefore c_1 = \cdots = c_m = 0$.

আর $c_{m+1} = 0$ দেখানোটা আগের মতই--

Hence, as in case I, $c_{m+1} = 0$.
So the proof is complete by the principle of mathematical induction.

[Q.E.D]

যদি A একটা Hermitian বা real, symmetric matrix হয়, তবে এই theorem-টাকে আরও একধাপ এগিয়ে নিয়ে যাওয়া যায়। তখন $\vec{v}_1, ..., \vec{v}_k$-গুলো linearly independent তো থাকবেই, সেই সঙ্গে আরো বাড়তি একটা শর্তও পালন করবে--

$$\forall i \neq j \quad \vec{v}_i^* \vec{v}_j = 0.$$

সেটাই এবার দেখাব।

> **THEOREM**
>
> Let A be a Hermitian or real, symmetric matrix. Let $\vec{v}_1, \vec{v}_2$ be two eigenvectors of A corresponding to eigenvalues $\lambda_1 \neq \lambda_2$. Then $\vec{v}_1^* \vec{v}_2 = 0$.

Proof: আমরা জানি যে Hermitian বা real, symmetric matrix-দের ক্ষেত্রে eigenvalue-গুলো real হয়।

> $\because A$ is Hermitian or real symmetric, $\therefore \lambda_1, \lambda_2 \in \mathbb{R}$.

এবার একটা কৌশল করব। আমাদের দেখাতে হবে $\vec{v}_1^* \vec{v}_2 = 0$. এর জন্য আমরা প্রথমে দেখাব যে $\lambda_1 \vec{v}_1^* \vec{v}_2 = \lambda_2 \vec{v}_1^* \vec{v}_2$. তাহলেই প্রমাণটা হয়ে যাবে, কারণ $\lambda_1 \neq \lambda_2$. এটা বুঝলে তো?
এবার তবে প্রমাণটা করি। কায়দাটা হল $\vec{v}_1^* A \vec{v}_2$-কে দুভাবে দলাইমলাই করে দুটো আলাদা রূপ দেওয়া। সে দুটো রূপ যে আসলে সমান, সেটা থেকেই প্রমাণটা হয়ে যাবে।

> Now $\vec{v}_1^* A \vec{v}_2 = \vec{v}_1^* (\lambda_2 \vec{v}_2) = \lambda_2 \vec{v}_1^* \vec{v}_2$.
> Again,
>
> $$\begin{aligned} \vec{v}_1^* A \vec{v}_2 &= (A\vec{v}_1)^* \vec{v}_2 \quad [\because A^* = A] \\ &= (\lambda_1 \vec{v}_1)^* \vec{v}_2 \\ &= \overline{\lambda}_1 \vec{v}_1^* \vec{v}_2 \\ &= \lambda_1 \vec{v}_1^* \vec{v}_2 \quad [\because \lambda_1 \in \mathbb{R}.]. \end{aligned}$$

এবার এই দুটো রূপকে মেলানো যাক।

> Thus
>
> $$\lambda_1 (\vec{v}_1^* \vec{v}_2) = \lambda_2 (\vec{v}_1^* \vec{v}_2).$$
>
> Since $\lambda_1 \neq \lambda_2$, we have $\vec{v}_1^* \vec{v}_2 = 0$, as required.

[Q.E.D]

Exercise 17: If all the eigenvalues of a complex $n \times n$ matrix are regular, then show that there is a basis $\{\vec{v}_1, ..., \vec{v}_n\}$ of $\mathbb{C}^n$ such that each $\vec{v}_i$ is an eigenvector of the matrix. ■

42.3 Invariant space

আমরা eigenspace-এর কথা শিখেছি। এর একটা ধর্ম হল এই যে, যদি A-র কোনো eigenspace-এর মধ্যে একটা vector নাও $\vec{v}$, তবে $A\vec{v}$-ও ওই eigenspace-এই থাকবে। এটা অবশ্য খুবই মামূলী শোনাতে পারে, কারণ eigenspace-টার জন্য eigenvalue-টা যদি λ হয়, তবে, $A\vec{v} = \lambda\vec{v}$ অবশ্যই $\vec{v}$-এর লাইন বরাবরই থাকবে। কিন্তু eigenspace ছাড়াও আরও এমন subspace সম্ভব (ধরো S), যাদের বেলায়ও এই ব্যাপারটা হয়, মানে $\vec{v} \in S$ হলে $A\vec{v} \in S$ হয়। নীচের সংজ্ঞাটা এদের নিয়েই।

DEFINITION: Invariance subspace

Let $T : V \to V$ be a linear transformation. Then a subspace $S \subseteq V$ is called an **invariant** subspace for T or a T-**invariant** subspace if

$$\forall \vec{v} \in S \quad T(\vec{v}) \in S.$$

এরকম subspace-এর সহজতম উদাহরণ হল $\{\vec{0}\}$ এবং V নিজে। এরা যেকোনো linear transformation-এর জন্যই invariant subspace. একটু আগেই দেখেছি যে, eigenspace-রাও হল invariant subspace. এবার নতুন কিছু invariant subspace দেখা যাক।

Example 21: ধরো A কোনো square matrix. তবে $\mathscr{C}(A^2)$ অবশ্যই একটা invariant subspace হবে। কেন?

SOLUTION: যদি $\vec{v} \in \mathscr{C}(A^2)$ হয়, তবে এমন $\vec{x}$ পাওয়া যাবে, যাতে $\vec{v} = A^2\vec{x}$ হয়। সুতরাং $A\vec{v} = A^3\vec{x} = A^2(A\vec{x}) \in \mathscr{C}(A^2)$. ■

এই অংকের কায়দাটা মাথায় রেখে নীচের অংকটা চেষ্টা করো।

Exercise 18: Let A be any square matrix, and let $p(x)$ be any polynomial. Show that $\mathscr{C}(p(A))$ is an invariant subspace for A. ■

Column space-এর বেলায় যা বলা গেল, সেটা কি row space বা nullspace-এর বেলাতেও বলা যাবে? উত্তরটা চেয়েছে নীচের অংকে।

Exercise 19: Let A be a square matrix, and let $p(x)$ be any polynomial. Then show that $\mathscr{N}(p(A))$ must be an invariant subspace for A, but $\mathscr{R}(p(A))$ need not be. ■

Example 22: Consider the linear transformation $T : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ defined by $T(A) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A$ for all $A \in M_{2\times 2}(\mathbb{R})$. Check whether the subspace $W = \{A \in M_{2\times 2}(\mathbb{R}) \ : \ A^t = A\}$ is T-invariant.[3] **(2020.CC12.3eii)**

SOLUTION:এরকম অংকে একটা উদাহরণ নিয়ে দেখলে সুবিধা হয়। W-এর সদস্যরা হল সব 2×2 সাইজের symmetric matrix. এরকম matrix-দের চেহারা হয়

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

এর উপর T লাগালে হবে--

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} b & c \\ a & b \end{bmatrix}.$$

এই matrix-টা symmetric নাই হতে পারে, যেমন যদি $a \neq c$ হয়। ব্যস্, আমরা বুঝে গেলাম যে W-টা মোটেই T-invariant নয়। এবার গুছিয়ে উত্তরটা লিখে ফেলি--

No, W is not T-invariant, because $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \in W$, but

$$T(A) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \notin W.$$

এই counter-example-টা হল $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ চেহারার, যেখানে $a \neq c$. ■

DAY 43
Transformations (part 1)

Eigenvalue বা eigenvector বার করাটা যে সবসময়ে নিতান্ত সহজ হয় না, তার পরিচয় তো পেয়েইছ। সুতরাং লোকে নানারকম কায়দা করে ঘুরপথে eigenvalue আর eigenvector বার করার চেষ্টা করে। একধরণের কায়দা হল যদি কোনো matrix-এর eigenvalue আর eigenvector জানা থাকে, সেটা ব্যবহার করে তার সঙ্গে সম্পর্কিত অন্য matrix-এর eigenvalue আর eigenvector বার করা। এরকম কিছু কায়দা এবার আমরা শিখব।

43.1 Transpose

যদি কোনো square matrix-এর transpose নাও, তবে তার eigenvalue বদলায় না। শুধু eigenvalue-ই নয়, তাদের algebraic বা geometric multiplicity-ও অপরিবর্তিত থাকে। তবে eigenvector-গুলো বেমালুম বদলে যেতে পারে। অনেক সময়ে এমন হয় যে, A-র eigenvalue-গুলো সরাসরি বার করা কঠিন হচ্ছে, অথচ A'-এর eigenvalue-গুলো হয় তো চোখে দেখেই বলে দেওয়া যায়। সেই যে Googleএর উদাহরণটা করেছিলাম, সেখানেই এই কায়দাটার একটা প্রয়োগ আছে। সেই প্রসঙ্গে এক্ষুণি আসছি, আগে যা বললাম অংকের ভাষায় লিখে প্রমাণ করে নিই।

THEOREM

If λ is an eigenvalue of a matrix A, then it is also an eigenvalue of A'. Also the algebraic and geometric multiplicities are the same in both cases.

Proof: Eigenvalue এবং তাদের algebraic multiplicity দুইই অপরিবর্তিত আছে দেখানোর জন্য সহজ পথ হল এটা দেখানো যে, characteristic polynomial-টাই A আর A'-র বেলায় একই।

Let the characteristic polynomial of A, A' be $p_A(\lambda), p_{A'}(\lambda)$, respectively.

$$\begin{aligned} \therefore p_{A'}(\lambda) &= |\lambda I - A'| \\ &= |(\lambda I - A)'| \quad [\because I' = I] \\ &= |\lambda I - A| \quad \left[\because \text{transpose does not change determinant}\right] \\ &= p_A(\lambda). \end{aligned}$$

ব্যস্, polynomial দুটোই যদি সমান হল, তবে তাদের root আর তাদের multiplicity সমান না হয়ে যায় কোথায়?

> So the roots and multiplicities of the two polynomials match.
> Hence A and A' have the same eigenvalues with the same algebraic multiplicities.

এবার geometric multiplicity-গুলো সমান দেখানোর পালা--

> Let λ be an eigenvalue of A (and A'). Then the geometric multiplicity of λ as an eigenvalue of A is $\text{nullity}(\lambda I - A)$.
> Similarly, the geometric multiplicity w.r.t. A' is $\text{nullity}(\lambda I - A')$.

আমরা জানি যে transpose নিলে rank বদলায় না, আর rank-এর সঙ্গে nullity-র একটা সম্পর্ক আছে--

rank+nullity=number of columns.

এইবার সেটা কাজে আসবে--

> Let A be $n \times n$.
> Then
>
> $$\begin{aligned} \text{nullity}(\lambda I - A) &= n - r(\lambda I - A) \\ &= n - r((\lambda I - A)') \quad \left[\because \text{transpose does not change rank.}\right] \\ &= n - r(\lambda I - A') \\ &= \text{nullity}(A' - \lambda I), \end{aligned}$$
>
> as required.

[Q.E.D]

আমরা আমাদের Google-এর উদাহরণে জোর দিয়ে বলেছিলাম যে, ইন্টারনেটটা যেভাবেই আঁকো না কেন, matrix-টার একটা eigenvalue অবশ্যই 1 হতে বাধ্য। সেই জোরটা এসেছিল এই theorem-টা থেকেই। নীচের অংকে সেটাই তুমি দেখবে।

Exercise 20: ধরো একটা matrix দিলাম A, যার প্রত্যেকটা row sum-ই α, মানে প্রথম row-এর সবগুলোর entry-র যোগফল α, দ্বিতীয় row-এর entry-গুলোর যোগফলও তাই, এইরকম। দেখাও যে, α অবশ্যই A-র একটা eigenvalue হবে। এর জন্য একটা eigenvector দিতে পারো? এবার ধরো এমন একটা matrix দিলাম B, যার column sum-গুলো সবাই α. তাহলে B-র একটা eigenvalue দিতে পারো? এবার Google-এর উদাহরণে গিয়ে দেখো তো সবগুলো column sum-ই 1 হতে বাধ্য কিনা।

HINT:
খালি একটা কথা বলে দিই। তার পরে তুমি নিজের চেষ্টা করবে। যদি একটা vector থাকে $\vec{v}$, তবে তার entry-গুলোর যোগফলকে এইভাবে লেখা যায়--

$$\vec{v}'\vec{\mathbb{1}},$$

যেখানে $\vec{\mathbb{1}}$ এমন একটা vector, যার সবগুলো entry-ই হল 1. ■

43.2 Inverse

আমরা জানি যে একটা nonsingular matrix-এর eigenvalue-রা সবাই nonzero হয়। যদি matrix-টাকে invert করো, তবে তার eigenvalue-গুলোও উল্টে যাবে, মানে λ থাকলে $1/\lambda$ হয়ে যাবে। Algebraic আর geometric multiplicity অপরিবর্তিত থাকবে, eigenvector-রাও বদলাবে না।

THEOREM

Let A be a nonsingular matrix. Let λ be an eigenvalue and $\vec{v}$ a corresponding eigenvector of A. Then $1/\lambda$ will be an eigenvalue of A^{-1} and $\vec{v}$ will be a corresponding eigenvector of A^{-1}. Also

$$\text{am}(\lambda, A) = \text{am}\left(\tfrac{1}{\lambda}, A^{-1}\right) \text{ and } \text{gm}(\lambda, A) = \text{gm}\left(\tfrac{1}{\lambda}, A^{-1}\right).$$

Proof: আমরা কী করে এগোব বলে নিই। ধরো A-র eigenvalue-রা হল $\alpha_1, ..., \alpha_n$. তাহলে দেখাব যে, A^{-1}-এর eigenvalue-রা হবে $\frac{1}{\alpha_1}, ..., \frac{1}{\alpha_n}$. মানে A^{-1}-এর characteristic polynomial-টাকে দেখাব

$$p_{A^{-1}}(\lambda) = \left(\lambda - \tfrac{1}{\alpha_1}\right) \cdots \left(\lambda - \tfrac{1}{\alpha_n}\right).$$

তাহলে eigenvalue আর algebraic multiplicity সমান দেখানো--দুইই একঢিলে হয়ে যাবে।

> Characteristic polynomial of A^{-1} is
>
> $$\begin{aligned} p_{A^{-1}}(\lambda) &= |\lambda I - A^{-1}| \\ &= |A^{-1}(\lambda A - I)| \\ &= |A^{-1}| \cdot |\lambda A - I| \end{aligned}$$

আমাদের পৌঁছতে হবে A-র characteristic polynomial-এ, মানে $p_A(\lambda)$-তে। এখানে λ-টা A-র সঙ্গে লেগে আছে, ওটাকে সরানোর জন্য λ-টাকে 'কমন' নিলে সুবিধা হবে। কিন্তু তার জন্য কাজ করতে হবে $\lambda \neq 0$ নিয়ে। অতএব--

> $\therefore$ For $\lambda \neq 0$ we have
>
> $$\begin{aligned} p_{A^{-1}}(\lambda) &= |A^{-1}| \cdot \left|-\lambda\left(\tfrac{1}{\lambda} I - A\right)\right| \\ &= (-\lambda)^n |A^{-1}| \cdot \left|\tfrac{1}{\lambda} I - A\right| \\ &= (-\lambda)^n |A^{-1}| p_A\left(\tfrac{1}{\lambda}\right). \end{aligned}$$

এইবার $\alpha_1, ..., \alpha_n$-রা মঞ্চে প্রবেশ করবে--

> Let the eigenvalues of A be $\alpha_1, ..., \alpha_n$ (not necessarily distinct).
>
> $\because A$ is nonsingular, $\therefore \forall i \quad \alpha_i \neq 0$.

এখানে "not necessarily distinct" লিখে নিয়েছি, যাতে algebraic multiplicity-গুলো হিসেবে এসে যায়।

Then $p_A(\lambda) = (\lambda - \alpha_1)\cdots(\lambda - \alpha_n)$.
So, for $\lambda \neq 0$,

$$\begin{aligned} p_{A^{-1}}(\lambda) &= (-\lambda)^n |A^{-1}| p_A\left(\tfrac{1}{\lambda}\right) \\ &= (-\lambda)^n |A^{-1}| \left(\tfrac{1}{\lambda} - \alpha_1\right)\cdots\left(\tfrac{1}{\lambda} - \alpha_n\right) \end{aligned}$$

এবার ওই $(-\lambda)^n$-এর প্যাকেটটা খুলে
প্রতিটা $\left(\frac{1}{\lambda} - \alpha_i\right)$-কে একটা করে $-\lambda$ উপহার দেব--

$$= |A^{-1}| (\alpha_1\lambda - 1)\cdots(\alpha_n\lambda - 1)$$

এবার α_i-দের 'কমন' নেব--

$$\begin{aligned} &= \frac{\alpha_1\cdots\alpha_n}{|A|}\left(\lambda - \tfrac{1}{\alpha_1}\right)\cdots\left(\lambda - \tfrac{1}{\alpha_n}\right) \\ &= \left(\lambda - \tfrac{1}{\alpha_1}\right)\cdots\left(\lambda - \tfrac{1}{\alpha_n}\right) \quad [\because |A| = \alpha_1\cdots\alpha_n]. \end{aligned}$$

এখানে $|A| = \alpha_1 \cdots \alpha_n$ হল, কারণ eigenvalue-দের গুণ করলে determinant পাওয়া যায়।

Thus

$$\forall \lambda \neq 0 \quad p_{A^{-1}}(\lambda) = \left(\lambda - \tfrac{1}{\alpha_1}\right)\cdots\left(\lambda - \tfrac{1}{\alpha_n}\right).$$

এইবার ওই $\lambda \neq 0$-র ঝামেলাটা দূর করতে হবে। সেটা খুবই সহজ কাজ, যুক্তিটা জানা না থাকলে ভালো করে বুঝে নাও, এটা মাঝে মাঝেই কাজে লাগে। Fig 20 দেখো। এখানে একটা continuous function-এর গ্রাফ এঁকেছি, খালি $f(0)$-র বিন্দুটা বাদ রেখেছি। সেই বিন্দুটা তোমাকে আঁকতে হবে। নিশ্চয়ই কোনো সন্দেহ নেই বিন্দুটা কোথায় বসবে। এ থেকেই বুঝতে পারছ যে, একটা continuous function যদি 0 ছাড়া বাকি সব জায়গায় জানা থাকে, তবে 0-তেও তার value নির্ধারিত হয়ে যায়। অংকের ভাষায় বললে কারণটা হল $f(0) = \lim_{x\to 0} f(x)$, আর $\lim_{x\to 0} f(x)$ বার করার জন্য $x \neq 0$-র জন্য $f(x)$ জানাই যথেষ্ট। এবার আমাদের অংকটায় ফিরে আসি। এখানে অবশ্য কাজ হচ্ছে $\mathbb{C}$-এর উপরে)$\mathbb{R}$-এর উপরে নয়)। কিন্তু limit-এর যুক্তিটা এখানেও প্রযোজ্য। দু দিকেই দুটো polynomial, তাই continuous function, এবং ওরা যাবতীয় $\lambda \neq 0$-র জন্য সমান। সুতরাং এই limit-এর যুক্তিতে ওরা সর্বত্রই সমান হতে বাধ্য।
এই যুক্তিটা যদি আগে না দেখে থাকো, তাই এত ব্যাখ্যা করে লিখলাম। কিন্তু অংকের জগতে এটা এতই সুপরিচিত জিনিস যে, প্রমাণের মধ্যে এক কথায় সেরে দিলেই চলে--

Since both sides are continuous, they must match at $\lambda = 0$ also, ie,

$$\forall \lambda \in \mathbb{C} \quad p_{A^{-1}}(\lambda) = \left(\lambda - \frac{1}{\alpha_1}\right)\cdots\left(\lambda - \frac{1}{\alpha_n}\right).$$

Fig 20

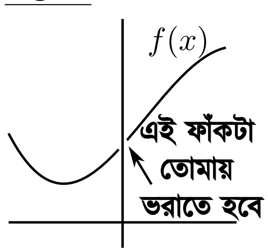
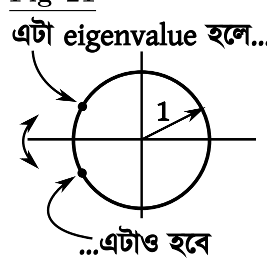


Fig 21



So the eigenvalues of A^{-1} are $\frac{1}{\alpha_i}$'s with algebraic multiplicities equal to the algebraic multiplicities of α_i's for A.

এবার eigenvector-দের নিয়ে পড়া যাক--

Also $\vec{v} \neq \vec{0}$ is an eigenvector of A for eigenvalue α_i if and only if

$$\begin{aligned} & A\vec{v} = \alpha_i \vec{v} \\ \iff\quad & A^{-1}A\vec{v} = \alpha_i A^{-1}\vec{v} \quad \left[\because A, A^{-1} \text{are nonsingular.}\right] \\ \iff\quad & \vec{v} = \alpha_i A^{-1}\vec{v} \\ \iff\quad & A^{-1}\vec{v} = \frac{1}{\alpha_i}\vec{v}, \end{aligned}$$

ie, if and only if $\vec{v}$ is an eigenvector of A^{-1} corresponding to eigenvalue $\frac{1}{\alpha_i}$.
So the eigenspaces of A and A^{-1} are identical, and hence the geometric multiplicities are the same, as required.

[Q.E.D]

আমরা এতক্ষণ যা যা শিখেছি তারই কয়েকটাকে মেলালেই নীচের অংকটা হবে।

Example 23: If λ is an eigenvalue of a real orthogonal matrix A, prove that $\frac{1}{\lambda}$ is also an eigenvalue of A.

SOLUTION:

$A_{n\times n}$ is real, orthogonal, and hence nonsingular. So 0 is not an eigenvalue.
$\therefore \lambda \neq 0.$
Also $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}. $\because A$ is orthogonal, $A^{-1} = A'$.
Thus $\frac{1}{\lambda}$ is an eigenvalue of A'.
Also the eigenvalues of A and A' are the same.
So $\frac{1}{\lambda}$ is an eigenvalue of A, as required.

■

এই অংকটাকে ছবি দিয়ে ভাবা যায়। আমরা আগেই বলেছি যে, কোনো real orthogonal matrix-এর eigenvalue-দের modulus হয় 1. তার মানে complex plane-এ eigenvalue-রা সবাই থাকে unit circle-এর উপরে (পরিধিটার উপরে, ভিতরে বলছি না)। এটা আমরা আগেই জানতাম। এখন এই অংক থেকে শিখলাম যে, পরিধির উপরে কোনো বিন্দু যদি eigenvalue হয়, তবে real axis বরাবর তার প্রতিফলনটাও একটা eigenvalue হতে বাধ্য (Fig 21), কারণ unit circle-এর উপরে λ আর $\frac{1}{\lambda}$ হয় পরস্পরের প্রতিফলন।

43.3 যোগ, বিয়োগ, গুণ

A-র একটা eigenvalue যদি λ হয়, আর B-এর একটা eigenvalue হয় μ, তবে কি সব সময়েই $A+B$-এর একটা eigenvalue হবে $\lambda+\mu$? উঁহু, সেটা হলে জীবনটা বড্ড বেশী সহজ হয়ে যেত। নীচের counterexample-টা দেখো।

Example 24: যদি $A = B = diag(1,-1)$ নিই, তবে তার একটা eigenvalue হল $\lambda = 1$, আরেকটা $\mu = -1$. কিন্তু $\lambda+\mu = 1+(-1) = 0$ মোটেই $A+B$-র একটা eigenvalue নয়। ■

অবশ্য যদি eigenvector দুটো মিলে যায়, তবে আশা আছে। নীচের অংকে একটা উদাহরণ দেখাচ্ছি।

Example 25: ধরো A, B দুটো square matrix. তাদেরই একই eigenvector আছে $\vec{v}$. তবে A-র বেলায় eigenvalue-টা λ, আর B-এর বেলায় μ. তবে দেখাও যে, $\lambda+\mu$ অবশ্যই $A+B$-এর একটা eigenvalue হবে, এবং তার জন্য একটা eigenvector হবে সেই $\vec{v}$.

SOLUTION:

$$(A+B)\vec{v} = A\vec{v} + B\vec{v} = \lambda\vec{v} + \mu\vec{v} = (\lambda+\mu)\vec{v}.$$

■

ঠিক একইরকম কথা গুণের ক্ষেত্রেও খাটে--

Exercise 21: Let A, B both have a common eigenvector $\vec{v}$ with corresponding eigenvalue λ for A and μ for B. Then show that AB must have $\vec{v}$ as an eigenvector corresponding to eigenvalue $\lambda\mu$. ■

43.4 AB আর BA

A আর B-এর eigenvalue জানা থাকলেও AB-র eigenvalue-দের সম্বন্ধে খুব কিছু সিদ্ধান্ত করা যায় না। কিন্তু AB-র eigenvalue-দের ভিত্তিতে BA-র eigenvalue-গুলো চমৎকার বার করে দেওয়া যায়। এবং এর জন্য A আর B-কে আলাদা করে square হবারও দরকার পড়ে না, যেমন A যদি $m \times n$ আর B যদি $n \times m$ হয়, তবে AB আর BA দুজনেই square হবে। এদের nonzero eigenvalue-রা সমান হতে বাধ্য। সেটা আমরা দুভাবে দেখাব। দুটো প্রমাণে দুটো আলাদা বোনাস আছে। প্রথম প্রমাণটা খুব সহজ, এবং তাতে eigenvalue-র সাথে সাথে eigenvector বার করারও একটা পথ পাওয়া যায়। আর দ্বিতীয় প্রমাণটায় একটা ছোট্টো কৌশল লাগলেও বাড়তি পাওনা হল এই যে, এখানে nonzero eigenvalue-দের algebraic multiplicity-গুলোকেও সমান দেখানো যায়।

> THEOREM
>
> Let $A_{m\times n}$ and $B_{n\times m}$ be any two matrices. Then AB and BA have the same *nonzero* eigenvalues. Also, if $\vec{v}$ is an eigenvector of AB for any eigenvalue $\lambda \neq 0$ then $B\vec{v}$ will be an eigenvector of BA for the same eigenvalue λ.

Proof:

> We have $AB\vec{v} = \lambda\vec{v}$.

Multiplying by B from left, $BAB\vec{v} = \lambda B\vec{v}$ or

$$BA(B\vec{v}) = \lambda(B\vec{v}).$$

এর চেয়ে সহজ যুক্তি আর হয় না। অবশ্য প্রমাণটা একটু বাকি আছে, দেখাতে হবে যে $B\vec{v} \neq \vec{0}$.

Also $B\vec{v} \neq \vec{0}$.

⟦ Because:

Otherwise $\lambda\vec{v} = AB\vec{v} = \vec{0}$ and so $\lambda\vec{v} = \vec{0}$.
But $\lambda \neq 0$, hence $\vec{v} = \vec{0} (\Rightarrow\Leftarrow)$.

⟧

এখানে contradiction[6] হল কারণ $\vec{v}$ যেহেতু একটা eigenvector বলা আছে, তাই $\vec{v} \neq \vec{0}$ হতে হবে।

$\therefore B\vec{v}$ is an eigenvector of BA, as required.

[Q.E.D]

এবার একই জিনিস আরেকভাবে দেখব, যাতে এটাও প্রমাণ হয়ে যাবে যে, AB আর BA-র nonzero eigenvalue-রা সমান তো বটেই, ওদের algebraic multiplicity-গুলোও দুইক্ষেত্রে সমান।

THEOREM

Let $A_{m\times n}$ and $B_{n\times m}$ be any two matrices with $m \leq n$. Let $p_{AB}(\lambda)$ and $p_{BA}(\lambda)$ be the characteristic polynomials of AB and BA, respectively. Then

$$\forall \lambda \in \mathbb{C} \quad p_{BA}(\lambda) = \lambda^{n-m} p_{AB}(\lambda).$$

Hence AB and BA have the same nonzero eigenvalues with the same algebraic multiplicities.

Proof: আমরা প্রথমে $\lambda \neq 0$-র জন্য দেখাব--

Take any $\lambda \in \mathbb{C},\ \lambda \neq 0$.

এইবারই কৌশলটা, যেটা বলে না দিলে ভেবে বার করা কঠিন--

Consider the $(m+n) \times (m+n)$ block matrix

$$\begin{bmatrix} I_m & A \\ B & \lambda I_n \end{bmatrix}.$$

এরকম block matrix-এর determinant বার করার দুটো ফর্মুলা ছিল, মনে আছে? মনে না থাকলে, এখন মনে করিয়ে দিচ্ছি--

[6] ওই $\Rightarrow\Leftarrow$ হল contradiction-এর চিহ্ন।

> We know that for any block matrix
>
> $$\begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$
>
> the determinant can be written as
>
> $$|P| \cdot |S - RP^{-1}Q| \text{ if } |P| \neq 0,$$
>
> and
>
> $$|S| \cdot |P - QS^{-1}R| \text{ if } |S| \neq 0.$$
>
> In our case, $|I_m| = 1 \neq 0$ and $|\lambda I_n| = \lambda^n \neq 0,$ since $\lambda \neq 0.$

সুতরাং আমাদের ক্ষেত্রে দুটো ফর্মুলাই প্রযোজ্য। দুটো ফর্মুলা থেকেই একই determinant আসা উচিত, অতএব--

> So expanding the determinant in two ways we get
>
> $$|I_m| \cdot |\lambda I_n - BI_m^{-1}A| = |\lambda I_n| \cdot |I_m - A(\lambda I_n)^{-1}B|,$$

মনে রেখো যে $|\lambda I_n| = \lambda^n$ এবং $(\lambda I_n)^{-1} = \frac{1}{\lambda} I_n$. সুতরাং--

> or
>
> $$\begin{aligned} |\lambda I_n - BA| &= \lambda^n \left| I_m - \tfrac{1}{\lambda} AB \right| \\ &= \lambda^n \left| \tfrac{1}{\lambda} (\lambda I_m - AB) \right| \\ &= \lambda^{n-m} |\lambda I_m - AB|. \end{aligned}$$
>
> $\therefore \forall \lambda \neq 0 \quad p_{BA}(\lambda) = \lambda^{n-m} p_{AB}(\lambda).$

এবার $\lambda = 0$ হলে কী হয় সেটা দেখতে হবে। অবশ্য দেখার তেমন কিছু নেই, দুটো continuous function যদি 0 ছাড়া বাকি সর্বত্র সমান হয়, তবে 0-তেও সমান না হয়ে যায় না। কায়দাটা একটু আগেও করেছি। তাই সংক্ষেপে লিখে দিই--

> Since both LHS and RHS are continuous functions, and they match for all nonzero values, hence they must match everywhere. Thus
>
> $$\forall \lambda \in \mathbb{C} \quad p_{BA}(\lambda) = \lambda^{n-m} p_{AB}(\lambda).$$
>
> as required.

[Q.E.D]

এই theorem-টা থেকে nonzero eigenvalue-এর algebraic multiplicity সমান দেখানো যাচ্ছে। একই theorem থেকে কিন্তু zero eigenvalue-দের algebraic multiplicity নিয়েও গুরুত্বপূর্ণ তথ্য পাওয়া যায়। সেটা বুঝতে পারলেই নীচের অংকটা হয়ে যাবে।

Exercise 22: দুটো matrix আছে $A_{5\times 7}$ আর $B_{7\times 5}$. বলা আছে যে AB-র জন্য 0 একটা eigenvalue, যার algebraic multiplicity হল 1. বল তো BA-র জন্য কি 0 একটা eigenvalue হবে? এবং যদি হয়, তবে তার

algebraic multiplicity কত হবে? ■

DAY 44
Transformations (part 2)

44.1 Polynomials

ধরো $p(x)$ কোনো একটা polynomial. যেমন হয়তো $p(x) = x^2 - 5x + 6$. এখানে x-এর জায়গায় আমরা যেকোনো scalar (মানে আমাদের ক্ষেত্রে complex সংখ্যা) বসাতে পারি। ইচ্ছে করলে কিন্তু একটা square matrix-ও বসানো যায়। যেমন, A যদি একটা square matrix হয়, তবে $p(A) = A^2 - 5A + 6I$ হবে। লক্ষ্য কর $p(x)$-এ যেখানে খালি 6 ছিল, $p(A)$-র বেলায় সেখানে $6I$ এসেছে।

সাধারণভাবে polynomial-দের যেমন যোগ বিয়োগ গুণ করা যায়, x-এর জায়গায় matrix থাকলেও সে সবই অক্ষুণ্ণ থাকে। যেমন, আমাদের উদাহরণে factorise (উৎপাদকে বিশ্লেষণ) করা যায়--

$$p(x) = (x-2)(x-3),$$

তাই

$$p(A) = (A-2I)(A-3I).$$

সবই একইরকম, খালি constant term-গুলোর সাথে I লেগে থাকবে।

যদি A-র eigenvalue-গুলো জানা থাকে, তবে $p(A)$-র eigenvalue-গুলোও অনায়াসে বার করে দেওয়া যায়। একটা অতি সহজ উদাহরণ দিয়ে শুরু করি।

Example 26: ধরো A-র একটা eigenvalue হল λ, এবং তার জন্য একটা eigenvector হল $\vec{v}$. তাহলে $2A$-র একটা eigenvalue, eigenvector বার কর দেখি।

SOLUTION: আমাদের বলা আছে যে, $\vec{v} \neq \vec{0}$ এবং $A\vec{v} = \lambda\vec{v}$. এর দুইপাশকে 2 দিয়ে গুণ করে দাও--

$$2A\vec{v} = 2\lambda\vec{v}.$$

সুতরাং দেখাই যাচ্ছে যে 2λ হল $2A$-র একটা eigenvalue এবং তার জন্য $\vec{v}$ হল একটা eigenvector. ■

এখানে 2-এর জায়গায় অন্য যেকোনো সংখ্যা নিলেও চলত। তার মানে একটা matrix-কে কোনো সংখ্যা দিয়ে গুণ করলে তার eigenvalue-গুলোও সেই সংখ্যা দিয়ে গুণ হয়ে যায়, কিন্তু eigenvector-গুলো একই থাকে।

এবার আরেক ধাপ এগোই।

Example 27: If λ is an eigenvalue of an $n \times n$ matrix A, then prove that λ^2 is an eigenvalue of A^2.[2] **(2022.CC6.3gii)**

SOLUTION:

> We have $\vec{v} \neq \vec{0}$ such that $A\vec{v} = \lambda\vec{v}$.

এর দুপাশকে বাঁদিক থেকে A দিয়ে গুণ করে দাও--

> Hence
> $$A^2\vec{v} = A(A\vec{v}) = A\lambda\vec{v} = \lambda A\vec{v} = \lambda^2\vec{v}.$$

অর্থাৎ λ^2 হল একটা eigenvalue, আর তার জন্য একটা eigenvector হল $\vec{v}$.

> Thus, λ^2 is an eigenvalue of A^2 (corresponding to eigenvector $\vec{v} \neq \vec{0}$).

■

এই অংকটা বুঝে থাকলে নীচের অংকটাও একইভাবে করতে পারা উচিত।

Exercise 23: Let A have an eigenvalue λ with corresponding eigenvector $\vec{v}$. Then show that A^n has an eigenvalue λ^n with corresponding eigenvector $\vec{v}$, where $n \in \mathbb{N}$. ■

এবার আরও একধাপ এগোব।

Example 28: যদি A-র একটা eigenvalue হয় λ, আর তার জন্য $\vec{v}$ হয় একটা eigenvector, তবে $A - 5A^2$-এর একটা eigenvalue বার কর। এবং তার জন্য একটা eigenvector কী হবে বলো।
SOLUTION: বলা আছে $\vec{v} \neq \vec{0}$ এবং $A\vec{v} = \lambda\vec{v}$. তা থেকে আমরা A দিয়ে গুণ করে পেতে পারছি $A^2\vec{v} = \lambda^2\vec{v}$. সুতরাং $(A - 5A^2)\vec{v} = (\lambda - 5\lambda^2)\vec{v}$. অতএব $(\lambda - 5\lambda^2)$ হল $A - 5A^2$-এর একটা eigenvalue, এবং তার জন্য একটা eigenvector হল সেই $\vec{v}$. ■

সুতরাং আন্দাজ করতে পারছ যে, A-র উপরে কোনো polynomial লাগালে eigenvalue-দের ঘাড়েও সেই একই polynomial এসে চাপে। নীচের theorem-টার বক্তব্যও সেটাই, সঙ্গে eigenvector নিয়েও কিছু কথা আছে।

THEOREM

Let A be any square matrix, and let $p(x)$ be any polynomial.

1. If $\vec{v}$ is an eigenvector of A for eigenvalue λ, then $\vec{v}$ is also an eigenvector of $p(A)$ for eigenvalue $p(\lambda)$.
2. If μ is an eigenvalue of $p(A)$ then μ must be of the form $p(\lambda)$ for some eigenvalue λ of A.

Proof: এখানে (1) আর (2) যেন কতকটা পরস্পরের বিপরীত। যদি A থেকে $p(A)$-তে যেতে চাও তবে (1) কাজে দেবে। আবার $p(A)$ থেকে A-তে ফিরতে চাইলে কাজে লাগবে (2). তবে একটা জিনিস লক্ষ্য কর যে, (1)-এ eigenvector-দের কথা বলা আছে, (2)-তে কিন্তু নেই। কেন নেই, সে নিয়ে কিছু কথা বলব প্রমাণটা হয়ে গেলে।

> (1) Let $p(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$.
> We have $\vec{v} \neq \vec{0}$ and $A\vec{v} = \lambda\vec{v}$. Multiplying by A from left repeatedly,
> $$\forall i \in \mathbb{N} \quad A^i\vec{v} = \lambda^i\vec{v}.$$

So

$$\begin{aligned} p(A)\vec{v} &= a_k A^k \vec{v} + \cdots + a_1 A\vec{v} + a_0 I\vec{v} \\ &= a_k \lambda^k \vec{v} + \cdots + a_1 \lambda \vec{v} + a_0 \vec{v} \\ &= (a_k \lambda^k + \cdots + a_1 \lambda + a_0)\vec{v} \\ &= p(\lambda)\vec{v}. \end{aligned}$$

So $\vec{v} \neq \vec{0}$ is an eigenvector of $p(A)$ corresponding to eigenvalue $p(\lambda)$.

এবার অন্য দিকটা--

(2) If $p(x) \equiv 0$, then the result is obvious.

এটা আবার কেন লিখলাম বলি। আমরা $p(x)$-কে এক্ষুণি linear factor-এ ভেঙে লিখব। এরকম factor-এর সংখ্যা হবে $p(x)$-এর degree-র সমান, যেমন $p(x) = x^2 + 1$ হলে degree হল 2, আর factor-ও দুটো, কারণ $p(x) = (x - i)(x + i)$. এই কথাটা খাটে না খালি একটা polynomial-এর বেলায়, সে হল zero polynomial, কারণ সেটাই একমাত্র polynomial যার কোনো finite degree হয় না (কেউ কেউ ধরেন যে, zero polynomial-এর degree বলে কিছু হয়ই না, কেউ কেউ বলেন degree-টা $-\infty$)। সেই কারণেই $p(x) \equiv 0$ কেসটাকে আলাদা করে লিখে নিয়েছি। এক্ষেত্রে $p(A) = O$, যার সবগুলো eigenvalue-ই 0.
এবার বাকি $p(x)$-দের নিয়ে কাজ করি--

Otherwise, let μ be any eigenvalue of $p(A)$. Let $\lambda_1, ..., \lambda_k$ be all the roots (counted according to multiplicities) of the polynomial $\mu - p(x)$. Then

$$\mu - p(x) = a_k(\lambda_1 - x) \cdots (\lambda_k - x),$$

where $a_k \neq 0$.

এইটা যদি গুলিয়ে যেতে চায়, তবে একটা উদাহরণ ভেবে নাও। ধরো $p(x) = 2x^2 + 5x + 1$ আর $\mu = -2$. তবে $\mu - p(x) = -2x^2 - 5x - 3 = \cdots = -2\left(x + \frac{3}{2}\right)(x+1) = -2\left(-\frac{3}{2} - x\right)(-1 - x)$. আমরা সাধারণতঃ polynomial-এর factor-দের $(x - \lambda_i)$ আকারে লিখে থাকি। এখানে যে $(\lambda_i - x)$ লিখলাম, তার কারণ আমরা এক্ষুণি characteristic polynomial ব্যবহার করব, যার সংজ্ঞা হল $|\lambda I - A|$, যেখানে λ-টা শুরুতে আছে।
আমরা $\mu - p(x)$-এর যে factorisation-টা পেলাম, তাতে x-এর জায়গায় A বসিয়ে দিলেও একই জিনিস লেখা যায় (খালি constant term-গুলোর সঙ্গে I জুড়ে দিতে হবে)।

So

$$\mu I - p(A) = a_k(\lambda_1 I - A) \cdots (\lambda_k I - A).$$

এবার দুদিকের determinant নেব--

$$\therefore |\mu I - p(A)| = |a_k|^n \times |\lambda_1 I - A| \cdots |\lambda_k I - A|,$$

where A is $n \times n$.

আমরা এখানে দুটো তথ্য ব্যবহার করেছি--এক, matrix-এর গুণফলের determinant হয় determinant-এর গুণফল। আর দুই, $|aA| = a^n|A|$.

> Since $|\mu I - p(A)| = 0$ and $a_k \neq 0$, hence at least one $|\lambda_i I - A|$ must be 0.

এখানে $|\mu I - p(A)| = 0$ হল, কারণ μ বলা আছে $p(A)$-র একটা eigenvalue.

> Choose $\lambda = \lambda_i$.
> Then $|\lambda I - A| = 0$, and so λ is an eigenvalue of A. Also $\mu - p(\lambda) = 0$, ie, $\mu = p(\lambda)$, as required.

[Q.E.D]

এই theorem-টা থেকে শিখলাম যে, A-র eigenvalue আর eigenvector থেকে $p(A)$-র eigenvalue আর eigenvector বার করা যায়। আবার $p(A)$-র eigenvalue থেকে A-এর eigenvalue-দের বিষয়েও কিছু সিদ্ধান্ত করা যায়। কিন্তু $p(A)$-র eigenvector থেকে A-র eigenvector-দের বিষয়ে কিছু বলা যায় না? দুঃখের কথা হল, তেমন কিছুই বলা যায় না। নীচের অংকে সেই দুঃখটাই ফেটে বেরিয়েছে!

Example 29: আমরা জানি যে, A-র eigenvector-রা A^2-এরও eigenvector হয়। কিন্তু A^2-এর eigenvector-রা কি A-র eigenvector হতে বাধ্য?

SOLUTION: নাঃ, ধরো

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

তবে $A^2 = I$, এবং A^2-এর একটা eigenvector হল $(1,1)$. কিন্তু এটা মোটেই A-র eigenvector হচ্ছে না। ■

এমনটা কেন হল? কারণ A-র দুরকম eigenvalue ছিল 1 আর -1, এরা দুজনেই পরিণত হয়েছে A^2-এর eigenvalue 1-এ। প্রত্যেকেই নিজেদের eigenvector-দের সঙ্গে করে এনেছে। ফলে A^2-এর একটা eigenvector হল $(1,0)$, যেটা ছিল A-র eigenvector 1-এর জন্য। আবার আরেকটা eigenvector হল $(0,1)$ যেটা এসেছিল -1 থেকে। সমস্যা হল এদের মিলিয়ে পাওয়া যায় $(1,0) + (0,1) = (1,1)$. এইবার যখন A^2 থেকে A-তে ফেরার চেষ্টা করছি, তখন $(1,1)$ কোন দিকে ফিরবে? শ্বশুরবাড়ি নাকি বাপেরবাড়ি?

নীচের অংকটায় একটা মৃদু আশার আলো আছে বটে, কিন্তু এর জন্য এমন বিদঘুটে শর্ত লাগছে যে ব্যাপারটা খুব কাজের নয়।

Exercise 24: ধরো $\vec{v}$ হল A^2-এর একটা eigenvector. এর জন্য eigenvalue-টা ধরো λ^2. যদি বলা থাকে যে λ হল A-র একটা eigenvalue, কিন্তু $-\lambda$ নয়, তবে দেখাও যে $\vec{v}$ অবশ্যই A-র একটা eigenvector হবে। এখানে শ্বশুরবাড়ি আর বাপের বাড়ির মধ্যে একটার পথ বন্ধ করে দিয়েছি। ■

আবার theorem-টার দিকে চোখ ফেরাই। ধরো A-র eigenvalue-রা হল

$$-3, 1, 3, 3, 5.$$

এখানে 3 সংখ্যাটা দুবার লেখার মানে 3-এর algebraic multiplicity হল 2. যে theorem-টা আমরা শিখলাম, তা থেকে বলতে পারি যে, A^2-এর eigenvalue-রা হবে $(-3)^2 = 3^2 = 9$, $1^2 = 1$ আর $5^2 = 25$. কিন্তু কোনটার algebraic multiplicity কত হবে সে বিষয়ে কিন্তু theorem-টা নীরব। সেই উত্তরটা আছে নীচের theorem-টায়। এটা বলছে যে, A^2-এর eigenvalue-রা হবে $(-3)^2, 1^2, 3^2, 3^2, 5^2$, মানে 9-এর algebraic multiplicity হল $1 + 2 = 3$ (একবার 9 আসছে -3 থেকে, আর দুইবার 3 থেকে), এবং বাকিদের algebraic multiplicity হল 1.

THEOREM

Let A be any square matrix, and let $p(x)$ be any polynomial. Let μ be any eigenvalue of $p(A)$. Let $\lambda_1, ..., \lambda_t$ be *all* the distinct eigenvalues of A such that $p(\lambda_i) = \mu$. Then

1. $\text{am}(\mu, p(A)) = \sum_{i=1}^t \text{am}(\lambda_i, A)$,
2. $\text{gm}(\mu, p(A)) \geq \sum_{i=1}^t \text{gm}(\lambda_i, A)$.

এর মধ্যে প্রথম অংশটার একটা অত্যন্ত সহজ প্রমাণ আছে, কিন্তু তার জন্য Schür decomposition বলে একটা জিনিস লাগবে, যেটা আমরা কালকে শিখব। তাই প্রমাণের সেই অংশটাও কালকের জন্যই তোলা থাক। দ্বিতীয় অংশটার প্রমাণ খুবই সহজ, সেটা এখনই করে ফেলি--ধরো λ_i-এর geometric multiplicity হল g_i. তাহলে প্রতিটা λ_i-এর জন্য g_i-খানা করে eigenvector পাবে, যারা linearly independent. মোট λ_i আছে t-খানা, সুতরাং এইভাবে আমরা $g_1 + \cdots + g_t$-খানা vector পাচ্ছি। এরা সকলেই $p(A)$-র eigenvector, এবং তখন এদের সবারই জন্য eigenvalue হল μ. এই vector-গুলো সবাই linearly independent (কেন বলো তো? প্রমাণটা সহজই, চেষ্টা করে দেখো! মনে রেখো যে, আলাদা আলাদা eigenvalue-র জন্য eigenvector-রা linearly independent হয়।) সুতরাং μ-এর জন্য $p(A)$-এর eigenspace-এ অন্ততঃ $\sum g_i$-খানা linearly independent vector আছে, অতএব তার dimension নিশ্চয়ই $\geq \sum g_i$. ব্যস্, প্রমাণ শেষ।

Exercise 25: এমন একটা উদাহরণ দিতে পারো, যেখানে $\text{gm}(\mu, p(A)) > \sum_{i=1}^t \text{gm}(\lambda_i, A)$ হয়?

HINT: একটু ধরিয়ে দিচ্ছি। $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ এবং A^2 নিয়ে চিন্তা করো।

■

এবার দুটো ছোট্টো কিন্তু গুরুত্বপূর্ণ প্রয়োগ--

Exercise 26: মনে কর A একটা idempotent matrix, মানে $A^2 = A$. দেখাও যে, এর eigenvalue-রা সকলেই হয় 0 নয়তো 1. ■

Exercise 27: একটা matrix আছে A. একে আমরা **nilpotent** বলি, যদি এমন কোনো $n \in \mathbb{N}$ পাওয়া যায়, যাতে $A^n = O$ হয়। বলতে পারো একটা nilpotent matrix-এর eigenvalue কী কী হতে পারে? ■

Exercise 28: If

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 2 & 0 \end{bmatrix},$$

show that A^2 cannot have imaginary characteristic roots.

■

44.2 Similarity

এইবার যেটা শিখতে চলেছি সেটা সহজ, কিন্তু তার গুরুত্ব অপরিসীম। খুঁটিনাটির মধ্যে যাবার আগে গুরুত্বর কথাটা বলে নিই। Fig 22-এ যে যন্ত্রটার ছবি দিয়েছি, সেটার সঙ্গে আমাদের চতুর্থ অধ্যায়েই মোলাকাত হয়েছিল। যদি $T : V \to V$

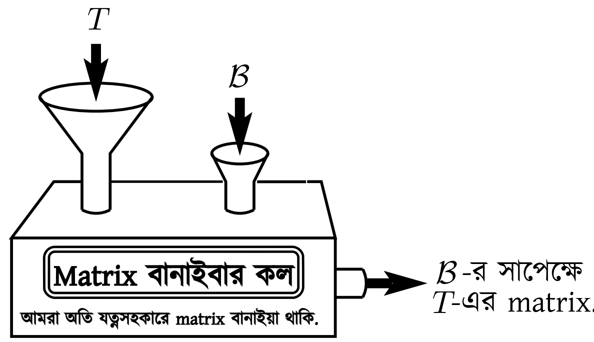


Fig 22

কোনো linear transformation হয়, তবে V-এর যেকোনো basis-এর সাপেক্ষে তার একটা matrix রূপ হয়। কী করে matrix-টা বার করতে হয়, সেটা যদি ভুলেও গিয়ে থাকো তো এখানে কোনো অসুবিধা নেই। খালি মনে রেখো যে, basis-টা বদলালে একই linear transformation-এর বিভিন্ন যেসব matrix পাওয়া যায়, তারা সবাই পরস্পরের সঙ্গে similar হতে বাধ্য। Similar মানে কী, মনে আছে তো?

> DEFINITION: **Similar matrix**
>
> Let A, B be two square matrices. B is called **similar** to A if there is a nonsingular matrix P such that $B = PAP^{-1}$.

কোনো কোনো বইতে দেখবে যে PAP^{-1}-এর জায়গায় $P^{-1}AP$ দেওয়া আছে। তাতে আসলে কিছুই বদলায় না, কারণ P-এর জায়গায় P^{-1} নিলেই ওই চেহারাটাও পেয়ে যাবে।

Similarity ব্যাপারটা অনেকটা fraction-দের কাটাকাটি করার মত। আমরা যেমন $\frac{50}{100}$ পেলে উপরনীচ থেকে 50 'কেটে নিয়ে' $\frac{1}{2}$ পাই। উপর থেকে 50 কেটে নেওয়া মানে $\frac{1}{50}$ দিয়ে গুণ করা, আর নীচ থেকে 50 কেটে নেওয়া মানে 50 দিয়ে গুণ করা। Similarity-র বেলাতেও তেমনি আমরা একটা কোনো square matrix নেব A, এবং তাকে কোনো একটা P এবং P^{-1} দিয়ে গুণ করব (যেখানে P কোনো একটা nonsingular matrix). যদি দুটো গুণই একদিক থেকে করি, তবে P আর P^{-1} তৎক্ষণাৎ কেটে যাবে, যেমন বাঁদিক থেকে করলে $P^{-1}PA = A$. তাতে মজা নেই। মজাটা হবে যদি P আর P^{-1} থাকে A-র দুদিকে, যেমন PAP^{-1} বা $P^{-1}AP$. যেহেতু matrix-এর গুণ commutative নয়, তাই এবার আর ওরা কেটে যেতে পারছে না।

এখানে কয়েকটা ব্যাপার লক্ষণীয়--

- যদি B হয় A-র সঙ্গে similar, তবে A-ও কিন্তু B-এর সাথে similar হতে বাধ্য। কারণ $B = PAP^{-1}$ হলে $A = P^{-1}BP$ হবে। সুতরাং আমরা বলতে পারি যে, similarity হল একটা symmetric relation. ঠিক Facebook-এ friend হওয়ার মত--তুমি আমার friend হলে আমিও তোমার friend. কিন্তু Facebook-এ follow করার মত নয়। সেখানে তুমি আমাকে follow করলেও আমি তোমাকে follow নাও করতে পারি।

- Similarity-র আরেকটা গুণ হল এটা transitive, মানে B যদি A-র সঙ্গে similar হয়, আর C হয় B-এর সঙ্গে similar, তবে C অবশ্যই A-র সঙ্গেও similar হতে বাধ্য। কারণ $B = PAP^{-1}$ আর $C = QBQ^{-1}$ হলে $C = QPAP^{-1}Q^{-1} = (QP)A(QP)^{-1}$ হবে।

- আরেকটা গুণ হল--similarity একটা reflexive relation-ও বটে, মানে যেকোনো A-ই নিজের সঙ্গে similar. কারণ P-এর জায়গায় I বসালেই দেখবে $A = IAI^{-1}$.

- তোমরা যারা equivalence relation-এর কথা পড়েছ, তারা জানো যে, একটা relation যদি একই সঙ্গে reflexive, symmetric আর transitive হয়, তবে তাকে বলে একটা equivalence relation. সুতরাং বুঝতেই পারছ যে, similarity হল একটা equivalence relation. যদি equivalence relation-এর কথা না পড়ে থাকো, তাতে বিচলিত হয়ো না। এই বইটা বোঝার জন্য সেটা গুরুত্বপূর্ণ নয়।

Relation-এর কচকচি তো অনেক হল, কিন্তু এটা তো এখনও বোঝা গেল না যে, A এবং PAP^{-1}-কে similar বা 'একইরকম' কেন বলা হচ্ছে, কী এমন আহামরি মিল আছে ওদের মধ্যে? Fig 22-এর দিকে তাকালেই বুঝবে যে, যে matrix-টা যন্ত্র থেকে বেরোচ্ছে, তার মধ্যে linear transformation এবং basis, এই দুয়ের ধর্ম মিলে আছে। Basis বদলালে matrix-টা হয়তো বদলাবে, কিন্তু তার মধ্যে linear transformation-ঘটিত যেসব ধর্ম ছিল, সেগুলো তো আর বদলাবে না। সেই ধর্মগুলোর সুবাদেই similar matrix-রা পরস্পরের সঙ্গে 'একরকম'। এরকম কিছু ধর্মের একটা তালিকা দিই। যদি A, B পরস্পরের সঙ্গে similar হয় (ধরো $B = PAP^{-1}$) তবে--

-- ওদের determinant সমান হবে, কারণ--

$$|B| = |PAP^{-1}| = |P| \cdot |A| \cdot |P^{-1}| = |P| \cdot |A| \cdot \frac{1}{|P|} = |A|.$$

-- ওদের trace সমান হবে, কারণ--

$$tr(B) = tr(PAP^{-1}) = tr(PP^{-1}A) = tr(A).$$

এখানে আমরা এই তথ্যটা ব্যবহার করেছি--যদি M, N এমন দুটো matrix হয়, যাতে MN এবং NM এই দুটো গুণফলই বার করা যায়, তবে $tr(MN) = tr(NM)$ হবে।

-- ওদের rank সমান হবে, কারণ nonsingular matrix দিয়ে গুণ করলে rank কখনও বদলায় না।

-- যদি A-টা nonsingular হয় (তবে B-ও তাই হবে, কারণ rank সমান), তাহলে A^{-1} আর B^{-1}-ও পরস্পরের সঙ্গে similar হবে, কারণ--

$$B^{-1} = \left(PAP^{-1}\right)^{-1} = \left(P^{-1}\right)^{-1} A^{-1} P^{-1} = PA^{-1}P^{-1}.$$

-- A' আর B' পরস্পরের সঙ্গে similar হবে, কারণ--

$$B' = \left(PAP^{-1}\right)' = \left(P^{-1}\right)' A' P' = \left(P'\right)^{-1} A' P' = QAQ^{-1},$$

যেখানে $Q = \left(P'\right)^{-1}$ আর তাই $Q^{-1} = P'$.

-- যেকোনো $n \in \mathbb{N}$-এর জন্য A^n আর B^n পরস্পরের সঙ্গে similar হবে, কারণ--

$$B^n = \underbrace{(PAP^{-1}) \times (PAP^{-1}) \times \cdots \times (PAP^{-1})}_{n \text{ বার}} = PA^nP^{-1}.$$

এখানে মাঝের $P^{-1}P$-গুলো সব কেটে যাচ্ছে।

-- একই যুক্তিতে, $q(x)$ যদি যেকোনো polynomial হয়, তবে $q(A)$ আর $q(B)$ হবে পরস্পরের সঙ্গে similar, কারণ--

$$q(B) = Pq(A)P^{-1}.$$

-- A আর B-এর যাবতীয় eigenvalue একই হবে। শুধু তাই নয়, তাদের algebraic এবং geometric multiplicity-গুলোও মিলে যাবে। এটা আমরা এক্ষুণি প্রমাণ করব।

-- A আর B-এর eigenvector-রা সমান নাও হতে পারে, কিন্তু ওদের মধ্যে একটা সহজ সম্পর্ক থাকবে--যদি $\vec{v}$ হয় A-র একটা eigenvector, তবে ওই একই eigenvalue-র জন্য B-এর একটা eigenvector হবে $P\vec{v}$. কারণটা এক লাইনে লিখে দিই--যদি $A\vec{v} = \lambda\vec{v}$ হয়, তবে

$$B(P\vec{v}) = PAP^{-1}P\vec{v} = PA\vec{v} = P\lambda\vec{v} = \lambda P\vec{v}.$$

ব্যাপারটাকে এইভাবেও ভাবতে পারো--এই লম্বা তালিকায় যেসব ধর্ম দিলাম, তারা সকলেই আসলে linear transformation-এর ধর্ম, তাই তুচ্ছ basis পরিবর্তন ওদের কোনো ক্ষতি করতে পারে না। যেমন, যদিও আমরা মুখে বলছি একটা matrix-এর eigenvalue, আসলে বলতে পারি একটা linear transformation-এর eigenvalue.

THEOREM

If A, B are similar matrices, then they have the same set of eigenvalues. Also, for each eigenvalue λ we have $\text{am}(\lambda, A) = \text{am}(\lambda, B)$ and $\text{gm}(\lambda, A) = \text{gm}(\lambda, B)$.

নীচের অংকে এই theorem-এর একটা অংশের প্রমাণ চেয়েছে।

Example 30: Prove that similar matrices have same eigenvalues.[2] **(2019)**

SOLUTION:

Let A and $B = PAP^{-1}$ be two similar matrices.
The characteristic polynomial of A is

$$p(\lambda) = |\lambda I - A|.$$

The characteristic polynomial of B is

$$\begin{aligned} q(\lambda) &= |\lambda I - PAP^{-1}| \\ &= |P(\lambda I - A)P^{-1}| \\ &= |P| \cdot |\lambda I - A| \cdot |P^{-1}| \\ &= |\lambda I - A| = p(\lambda). \end{aligned}$$

Thus both A and B have the same characteristic polynomial. Hence their eigenvalues (which are roots of the characteristic polynomial) must also be the same, as required.

■

লক্ষ্য কর যে, এখানে খালি eigenvalue-গুলো সমান দেখাতে বলেছিল। কিন্তু আমরা বিশল্যকরণীর বদলে একেবারে গন্ধমাদন এনে হাজির করেছি, পুরো characteristic polynomial-টাকেই সমান দেখিয়ে দিয়েছি! সুতরাং এতে এটাও প্রমাণ হয়ে গেল যে, algebraic multiplicity-গুলোও মিলতে বাধ্য।
আবার প্রায় একই অংক, কিন্তু ভাষাটা বদলেছে।

Exercise 29: If A is any square matrix and B is a nonsingular matrix of the same order, and if $C = BAB^{-1}$ and λ is a any scalar, then prove that

$$|C + \lambda I| = |A + \lambda I|.$$

HINT:
এখানে $|A+\lambda I|$ দেখে ঘাবড়ে যেও না। আমরা এতক্ষণ যে $|\lambda I - A|$ নিয়ে কাজ করছিলাম এটাও আসলে প্রায় তাই। সুতরাং প্রমাণটা একইরকম থাকবে। ■

Similar matrix-দের eigenvalue-রা যে এক হয়, সেটা দেখালাম। Algebraic multiplicity-রাও যে অভিন্ন হবে, সেটাও দেখানো হল। কিন্তু geometric multiplicity-রাও যে মিলবে, সে প্রমাণটা তো হল না! সে দায়িত্বটা তোমার উপরেই ছেড়ে দিয়েছি নীচের অংকটায়।

Exercise 30: ধরো $B = PAP^{-1}$ আর λ হল A-র (অতএব B-এরও) একটা eigenvalue. ধরো A-র জন্য λ-র eigenspace-এর একটা basis হল $\{\vec{v}_1, ..., \vec{v}_g\}$. (সুতরাং eigenspace-টার dimension, মানে λ-র geometric multiplicity দাঁড়াচ্ছে g). দেখাও যে $\{P\vec{v}_1, ..., P\vec{v}_g\}$ হল λ-র জন্য B-এর eigenspace-এর একটা basis. তা থেকেই বুঝতে পারছ $\text{gm}(\lambda, B) = g$.
HINT: $P\vec{v}_i$-রা যে eigenvector হবে, সে তো একটু আগেই দেখিয়েছি, similar matrix-দের eigenvector-দের সাদৃশ্য আলোচনার সময়ে। তোমায় দেখাতে হবে যে, এরা linearly independent এবং eigenspace-টার একটা generating set, অর্থাৎ $\vec{w}$ যদি λ-র জন্য B-এর কোনো eigenvector হয়, তবে $\vec{w}$-কে $P\vec{v}_i$-দের linear combination আকারে লিখে ফেলা যাবে। খুব কঠিন নয়, চেষ্টা করেই দেখো না! ■

DAY 45
জীবন সহজ করা

যদি তোমাকে বলি $\frac{36}{72} + \frac{50}{150}$ বার করতে, তবে তুমি নিশ্চয়ই প্রথমে fraction দুটোকে কাটাকাটি করে অংকটাকে $\frac{1}{2} + \frac{1}{3}$ বানিয়ে নেবে। কিন্তু যদি বলতাম $(36 + 72i) + (50 + 150i)$ বার করতে, তবে নিশ্চয়ই তুমি সেখানে একই কাটাকাটি করতে না! এই বোকা-বোকা কথাটার পিছনে একটা দরকারী সত্য লুকিয়ে আছে। আমরা কোনো অংক পেলেই সেটাকে যথাসম্ভব সহজ করে নেবার চেষ্টা করি, তার মূল নির্যাসটুকুকে অক্ষুণ্ণ রেখে। একটা fraction-এর ক্ষেত্রে মূল নির্যাস হল numerator আর denominator-এর ratio-টা, যেটা কাটাকাটি করলেও অপরিবর্তিত থাকে। তাই সেখানে দিব্যি কাটাকাটি করা চলে। কিন্তু complex সংখ্যার বেলায় real আর imaginary অংশদুটো আলাদা করে গুরুত্বপূর্ণ, কাটাকাটি করলে সেগুলো বদলে যায়। ঠিক একই ব্যাপার আমরা দেখব eigenvalue-র বেলাতেও। আমরা দেখেছি যে, দুটো matrix যদি similar হয়, তবে তাদের নানান ধর্ম একই হয়। তাই এই সব ধর্ম নিয়ে কাজ করার সময়ে যদি কোনো বিদঘুটে matrix আসে, তবে আমাদের চেষ্টা থাকে ওর সঙ্গে similar কোনো সহজ matrix বার করে রাখার, কারণ তাহলে ওই সহজ matrix-টাকে নিয়ে কাজ করলেই চলবে। সহজ matrix বলতে প্রথমেই মনে আসে diagonal আর triangular matrix-দের কথা, কারণ এদের determinant, eigenvalue ইত্যাদি বার করা খুবই সহজ। এদের মধ্যে diagonal-রা আবার একটু বেশী সহজ, এদের rank আর eigenvector-ও সহজে বার করা যায়। সুতরাং প্রশ্ন জাগে--যেকোনো matrix A-এর জন্যই কি এমন diagonal matrix বা triangular matrix পাওয়া যায়, যেটা A-র সঙ্গে similar? এই প্রশ্নের উত্তরই এবার আমরা শিখব।

45.1 Diagonalisation

Diagonal matrix-এর বেলায় এই প্রশ্নের উত্তর হল--না, এমন কোনো গ্যারান্টি নেই যে, সব square matrix-ই কোনো না কোনো diagonal matrix-এর সঙ্গে similar হবে। নীচের উদাহরণটা দেখলে সেটা সহজেই মালুম হবে।

Example 31: একটা triangular matrix নাও যার diagonal entry-গুলো সব শূন্য, কিন্তু অন্ততঃ একটা off-

diagonal entry আছে যেটা শূন্য নয়, যেমন ধরো এই matrix-টা--

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

দেখাও যে, এটা কোনো diagonal matrix-এর সঙ্গে similar হতে পারে না।

SOLUTION: এটা একটা triangular matrix, তাই এর eigenvalue-রা হল এর diagonal entry-গুলোই, মানে এক্ষেত্রে 0. এটা যদি কোনো diagonal matrix-এর সঙ্গে similar হয়, তবে সেই diagonal matrix-এর যাবতীয় diagonal entry-কেও একই কারণে শূন্য হতে হবে, মানে সেটা হবে O. কিন্তু তাহলে তো গোড়ার matrix-টাও O হয়ে যাচ্ছে, কারণ O-র সাথে similar matrix একমাত্র O নিজেই! কিন্তু আমাদের গোড়ার matrix-টা তো O ছিল না। তাই সেটা কোনো diagonal matrix-এর সাথে similar হতে পারে না। ■

আরো আলোচনার আগে একটা নতুন ভাষা শিখে রাখি। যদি কোনো matrix কোনো diagonal matrix-এর সঙ্গে similar হয়, তবে তাকে বলে **semisimple** বা **diagonalisable**.

> DEFINITION: **Semisimple/diagonalisable**
>
> A matrix is called **semisimple** or **diagonalisable** if it is similar to a diagonal matrix.

একটা matrix দেওয়া থাকলে কী করে বুঝবে সেটা diagonalisable কিনা? সেই কাজে eigenvalue আর eigenvector-রা খুব সাহায্য করবে। কী করে বলি।

ধরো A একটা diagonalisable matrix. তার মানে আমরা A-কে এইভাবে লিখতে পারি--

$$A = PDP^{-1},$$

যেখানে D হল diagonal, আর P হল nonsingular. ওই P^{-1}-টাকে বাঁদিকে নিয়ে আসি--

$$AP = PD.$$

যদি P-এর j-th column-টাকে $\vec{p}_j$ আর D-এর j-th column-টাকে $\vec{d}_j$ নাম দিই, তবে

$$A\vec{p}_j = P\vec{d}_j$$

পাব। এদিকে $\vec{d}_j$-র মধ্যে তো সব entry-ই শূন্য, খালি j-th entry-টাই বড়জোর nonzero, ওটাকে λ_j নাম দিলে $\vec{d}_j = \lambda_j\vec{e}_j$. এখানে $\vec{e}_j$ মানে সেই vector-টা যার সর্বত্রই 0, খালি j-th entry-টা হল 1. সুতরাং

$$A\vec{p}_j = \lambda_j P\vec{e}_j = \lambda_j\vec{p}_j.$$

এখানে আমরা এই তথ্যটা কাজে লাগিয়েছি যে, কোনো matrix-কে ডানদিক থেকে $\vec{e}_j$ দিয়ে গুণ করলে matrix-টার j-th column-টা পাওয়া যায়।

সুতরাং দেখা যাচ্ছে যে, P-এর column-গুলো সবাই A-র একেকটা eigenvector. এদিকে P আবার nonsingular-ও বটে, তার মানে ওর column-গুলো linearly independent-ও হবে। তার মানে A-র n-খানা linearly independent eigenvector পাওয়া যাচ্ছে। এর বিপরীত দিকটাও সত্যি, মানে A-র যদি n-খানা linearly independent eigenvector থাকে, তবে সেটা অবশ্যই diagonalisable হবে, কারণ ওই eigenvector-গুলোকে পাশাপাশি column হিসেবে দাঁড় করিয়ে দিলেই P পেয়ে যাব। এটাকে একটা theorem হিসেবে লিখে রাখা যাক--

THEOREM

A matrix $A_{n\times n}$ is diagonalisable if and only if A has n linearly independent eigenvectors.

গুলিয়ে যাচ্ছে? একটা উদাহরণ দেখে নেওয়া যাক।

Example 32: Find a nonsingular matrix P such that $P^{-1}AP$ is a diagonal matrix where

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$$

(2013,2008)

SOLUTION: যদি diagonal matrix-টাকে D নাম দিই তবে এমন P খুঁজছি যাতে $P^{-1}AP = D$, মানে $A = PDP^{-1}$ হয়। মানে A-কে diagonalise করার চেষ্টা করছি। এতক্ষণের আলোচনার ভিত্তিতে আমাদের কর্তব্য ঠিক করে নেওয়া যাক--

- প্রথমে eigenvalue-গুলো বার করব।
- তারপর প্রত্যেকটা eigenvalue-র জন্য eigenvector বার করব।
- এইভাবে তিনটে eigenvector যদি বার করে ফেলতে পারি, যারা linearly independent তো ব্যস্, তাদেরকে column-এর মত পাশাপাশি দাঁড় করে একটা matrix বানিয়ে নিলেই P পেয়ে যাব।

তাহলে eigenvalue বার করা দিয়ে শুরু করি--

> The characteristic polynomial is
>
> $$|\lambda I - A| = \lambda^3 - 7\lambda^2 + 11\lambda - 5 = (\lambda - 5)(\lambda - 1)^2.$$
>
> So the eigenvalues are 5 and 1.

এবার 5-এর জন্য eigenvector বার করতে হবে--

> The system $(5I - A)\vec{v} = \vec{0}$ is
>
> $$\begin{aligned} (5-2)x - 2y - z &= 0, \\ -x + (5-3)y - z &= 0, \\ -x - 2y + (5-2)z &= 0. \end{aligned}$$
>
> Solving we get $x = y = z$, and $x \in \mathbb{C}$ arbitrary.

এই সমাধানের ধাপগুলো আর দেখালাম না। ওটুকু তুমি রাফে করে নিও।

> So the eigenspace of A for eigenvalue 5 is
>
> $$\mathscr{N}(5I-A) = \{(x,x,x)\ :\ x\in\mathbb{C}\} = \text{span}\{(1,1,1)\}.$$

এর dimension হল 1, তাই--

> Thus the geometric multiplicity of 5 is 1, and an eigenvector of A corresponding to the eigenvalue 5 is $(1,1,1)$.

এবার অন্য eigenvalue-টা, মানে 1, নিয়ে পড়া যাক--

> The system $(I-A)\vec{v}=\vec{0}$ is
>
> $$\begin{aligned}(1-2)x-2y-z &= 0,\\ -x+(1-3)y-z &= 0,\\ -x-2y+(1-2)z &= 0.\end{aligned}$$

লক্ষ্য কর, তিনটে equation-ই আসলে একই জিনিস, $-x-2y-z=0$. সুতরাং এ থেকে খালি x,y,z-এর মধ্যে একজনকে বাকিদের দিয়ে প্রকাশ করা যাবে, ধরো z-কে প্রকাশ করলাম x,y দিয়ে। তাহলে পাব $z=-x-2y$. এর চেয়ে বেশী কিছু বলা যাচ্ছে না, অর্থাৎ--

> Solving we get $z=-x-2y$, and $x,y\in\mathbb{C}$ arbitrary.
>
> So the eigenspace for 1 is
>
> $$\begin{aligned}\mathscr{N}(I-A) &= \{(x,y,-x-2y)\ :\ y,z\in\mathbb{C}\}\\ &= \{x(1,0,-1)+y(0,1,-2)\ :\ x,y\in\mathbb{C}\}\\ &= \text{span}\{(1,0,-1),(0,1,-2)\}.\end{aligned}$$

এবার দেখাব যে $\{(1,0,-1),(0,1,-2)\}$ আসলে $\mathscr{N}(I-A)$-এর একটা basis. এটা যে একটা generating set সেটা দেখালামই। বাকি আছে খালি linearly independent দেখানো।

> Now $(1,0,-1)$ and $(0,1,-2)$ are linearly independent, because they are not multiples of each other.
>
> $\therefore$ Two linearly independent eigenvectors for 1 are $(0,1,0)$ and $(0,0,1)$.
>
> Since eigenvectors corresponding to distinct eigenvalues are linearly independent, so
>
> $$\{(1,1,1),(1,0,-1),(0,1,-2)\}$$
>
> is a linearly independent set of eigenvectors.

এইবার এদেরকে পাশাপাশি দাঁড় করিয়ে P বানালেই কাজ শেষ--

Thus we have

$$A\begin{bmatrix}1\\1\\1\end{bmatrix} = 5\times\begin{bmatrix}1\\1\\1\end{bmatrix},$$
$$A\begin{bmatrix}1\\0\\-1\end{bmatrix} = 1\times\begin{bmatrix}1\\0\\-1\end{bmatrix},$$
$$A\begin{bmatrix}0\\1\\-2\end{bmatrix} = 1\times\begin{bmatrix}0\\1\\-2\end{bmatrix}.$$

Combining

$$A\underbrace{\begin{bmatrix}1&1&0\\1&0&1\\1&-1&-2\end{bmatrix}}_{P}=\underbrace{\begin{bmatrix}1&1&0\\1&0&1\\1&-1&-2\end{bmatrix}}_{P}\underbrace{\begin{bmatrix}5&0&0\\0&1&0\\0&0&1\end{bmatrix}}_{D}.$$

So $AP = DP$, and P is a nonsingular matrix. Thus $P^{-1}AP = D$, a diagonal matrix, as required.

■

এবার তোমার জন্য একইরকম একটা অংক।

Exercise 31: Find a matrix P such that $P^{-1}AP$ is a diagonal matrix where

$$A=\begin{bmatrix}2&-1&1\\-1&2&-1\\1&-1&2\end{bmatrix}.$$

■

কোনো matrix-কে চোখে দেখেই কি বোঝা সম্ভব সেটা diagonalisable কিনা? দুঃখের কথা যে, উত্তর হল--নাঃ। কিন্তু এক ধরণের matrix আছে যারা diagonalisable হয়, এবং ওদেরকে চোখে দেখেই চেনা যায়। এরা হল real, symmetric matrix-রা[7]। এই ব্যাপারটা অংকের জগতে একটা গুরুত্বপূর্ণ theorem. আমরা এই বইতে পরে এটা গুছিয়ে লিখে প্রমাণ করব।
চট্ করে চোখে দেখেই যদিও diagonalisable-দের সবসময়ে সনাক্ত করা যায় না, কিন্তু যদি সবগুলো eigenvalue-র geometric multiplicity জানা থাকে, তা থেকেই অভ্রান্তভাবে বলে দেওয়া যায় matrix-টা diagonalisable হবে কিনা। একটা উদাহরণ দেখলেই বুঝবে।

Example 33: ধরো একটা 5×5 সাইজের matrix আছে A. তার দুটো distinct eigenvalue, যাদের geometric multiplicity হল 2 আর 3. বলতে পারো এই matrix-টা diagonalisable হবে কিনা?

[7]Complex Hermitian-দের বেলাতেও একই কথা বলা যায়।

SOLUTION: আমরা জানি যে, একটা 5×5 সাইজের matrix-এর diagonalisable হবার জন্য দরকার খালি 5-টা linearly independent eigenvector. ওদেরকে খাড়া করে পাশাপাশি দাঁড় করে দিলেই একটা nonsingular matrix পেয়ে যাব, P, যাতে $AP = PD$ হবে। এখানে একটা eigenvalue-র geometric multiplicity বলা আছে 2. তার জন্য দুটো eigenvector পাবে, যারা linearly independent. একইভাবে অন্য eigenvalue-টা থেকে আসবে তিনটে। আগের দুটো আর এই তিনটে মিলে linearly independent হবে, কারণ আলাদা আলাদা eigenvalue-দের eigenvector-রা linearly independent হয়। ব্যস্, সব মিলিয়ে $2 + 3 = 5$-টা linearly independent eigenvector পেয়ে গেলে। এখন আর diagonalisable হওয়া আটকায় কে? ■

এই অংকের যুক্তিটা থেকেই আসে নীচের theorem-টা--

> **THEOREM**
>
> A matrix $A_{n\times n}$ is diagonalisable if and only if the sum of the geometric multiplicities of all its distinct eigenvalues is n.

Geometric multiplicity-দের যোগফল n হওয়ার গল্প থেকে একটা জিনিস মনে পড়ে যায়-- algebraic multiplicity-দের যোগফল সর্বদাই n হয়। এদিকে কোনো eigenvalue-র geometric multiplicity কখনো তার algebraic multiplicity-কে টপকে যেতে পারে না। সুতরাং একটাও geometric multiplicity যদি algebraic multiplicity-র চেয়ে ছোটো হয়ে যায়, তবে geometric multiplicity-দের যোগফল আর n অব্দি উঠতে পারবে না। এই চিন্তা থেকেই আসে নীচের theorem-টা।

> **THEOREM**
>
> A matrix is diagonalisable if and only if all its eigenvalues are regular.

কোনো eigenvalue-এর algebraic multiplicity আর geometric multiplicity সমান হলে যে eigenvalue-টাকে regular বলে, সেটা মনে আছে নিশ্চয়ই?
এক্ষুণি যে অংকগুলো করলাম, সেখানে অধিকাংশ পরিশ্রমটাই যাচ্ছিল eigenvector-গুলো বার করতে, কারণ ওই eigenvector-গুলো দিয়েই P তৈরী হবে। কিন্তু যদি P-টা বার করার দরকার না থাকে, খালি diagonalise করা যায় কিনা সেটুকু দেখতে বলে, আর D বার করলেই হয়, তবে এই দুটো theorem-টা ব্যবহার করা বেশী সহজ। যেমন নীচের অংকটায়--

Example 34: Find all the eigenvalues of

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}.$$

Show that this matrix is diagonalisable. Write a diagonal matrix similar to this matrix.

SOLUTION:

> The characteristic polynomial is
>
> $$|\lambda I - A| = \lambda^3 - 12\lambda - 16 = (\lambda - 4)(\lambda + 2)^2.$$
>
> So the eigenvalues are 4 and -2 with algebraic multiplicities $1, 2$, respec-

> tively.

আমরা দেখাব যে, সবগুলো eigenvalue-ই regular.
আমরা জানি algebraic multiplicity $\geq$ geometric multiplicity হয়, এবং geometric multiplicity আবার সর্বদা ≥ 1 হয়। তাই algebraic multiplicity $= 1$ হলে বাধ্য হয়েই geometric multiplicity $= 1$ হতে হবে। মানে এইরকম eigenvalue-রা হবে regular.

> $\because$ Any eigenvalue with algebraic multiplicity 1 must be regular,
>
> $\therefore 4$ is a regular eigenvalue.

তাহলে 4-এর হিল্লে হয়ে গেল। এইবার -2-কে নিয়ে পড়ি।

> $$\mathrm{gm}(-2, A) = \mathrm{nullity}(-2I - A) = 3 - r(-2I - A).$$
>
> Now
> $$-2I - A = \begin{bmatrix} -2-1 & 3 & -3 \\ -3 & -2+5 & -3 \\ -6 & 6 & -2-4 \end{bmatrix} = \begin{bmatrix} -3 & 3 & -3 \\ -3 & 3 & -3 \\ -6 & 6 & -6 \end{bmatrix}.$$
>
> All the rows of this matrix are multiples of the first row (which is nonzero). So $r(-2I - A) = 1$.
>
> $\therefore \mathrm{nullity}(-2I - A) = 3 - 1 = 2 = \mathrm{am}(-2, A)$.
>
> Thus, all the eigenvalues of A are regular.

সুতরাং আমাদের theorem-টা লাগানো যাবে।

> Hence A is diagonalisable.
>
> A diagonal matrix similar to A is
> $$D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

এটা পেয়েছি eigenvalue-গুলোকে diagonal বরাবর লিখে। যেটার algebraic multiplicity যত, সেটা ততবার লিখেছি। এখানে eigenvalue-গুলো লেখার ক্রমটা গুরুত্বপূর্ণ নয়। ■

এই অংকটাতে diagonal matrix-টা বার করতে বলেছিল। তাই eigenvalue-গুলো লেগেছিল। নীচের অংকটায় সেটাও বার করতে হবে না, খালি diagonalisable দেখালেই হবে। তাই এখানে eigenvalue-গুলো বার না করেই কাজ সারার চেষ্টা করব।

Example 35: Show that the matrix

$$\begin{bmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

is diagonalisable.

SOLUTION:

> The characteristic polynomial is
>
> $$p(\lambda) = |\lambda I - A| = \cdots = \lambda^3 - 2\lambda^2 - \lambda + 6.$$

এইটা বার করার ধাপগুলো মামুলী বলে এখানে আর দেখালাম না।

কিন্তু এটার root বার করাটা মোটেই মামূলী মনে হচ্ছে না। ঠাণ্ডা মাথায় ভেবে নাও, root-গুলো বার করার সত্যিই কি খুব দরকার আছে? আমাদের তো D-টা বার করতে বলেনি। যদি কোনোভাবে eigenvalue-গুলোকে বার না করেই ওদের regular দেখিয়ে দিতে পারি, তবেই কেল্লা ফতে! এমনটা করা সহজ হবে যদি এটা দেখাতে পারি যে, eigenvalue-গুলো সবাই distinct. কারণ তাহলে ওদের algebraic multiplicity-গুলো সব 1 হবে, এবং সেক্ষেত্রে geometric multiplicity-গুলোও 1 হতে বাধ্য হবে, কারণ

$$1 \leq \text{geometric multiplicity} \leq \text{algebraic multiplicity}$$

হয়।

সুতরাং আমরা এটা দেখানোর একটা চেষ্টা করি যে, এই polynomial-টার সবগুলো root হবে distinct. এটা অবশ্যই eigenvalue-গুলো regular হবার চেয়ে একটা stronger condition. কিন্তু অনেক সময়ে stronger condition প্রমাণ করা সহজতর হয়।

কোনো polynomial-এর সবগুলো root আলাদা আলাদা কিনা সেটা বার করার একটা কায়দা আছে। তার জন্য root-গুলোকে ধরে ধরে বার করতে হয় না। যদি $p(x)$-এর কোনো দুটো root সমান হয়, তবে সেইটা $p'(x)$-এরও root হতে বাধ্য। এখানে $p'(x)$ মানে হল $p(x)$-এর derivative. একটা উদাহরণ দেখে নাও। ধরো $p(x) = (x-1)^2(x-2)$. তাহলে $p'(x) = 2(x-1)(x-2) + (x-1)^2$, যার থেকে একটা $(x-1)$ "কমন" নেওয়া যায়। এই যে "কমন" নেওয়া গেল, তার কারণ $p(x)$-এর মধ্যে $(x-1)$ ছিল দুইবার, তাই differentiate করার পরও একটা বেঁচে রয়েছে।

এ থেকে কোনো polynomial-এর সবগুলো root কখন distinct হবে, তার একটা শর্ত দেওয়া যায়--যখন $p(x)$ আর $p'(x)$-এর দুজনের কোনো common root থাকবে না। এখানে "common root" মানে এমন কোনো সংখ্যা, যেটা $p(x)$ আর $p'(x)$, দুজনেরই root. বা ঘুরিয়ে বললে, যখন $p(x)$ আর $p'(x)$-এর gcd-তে (মানে গসাগু-তে) কোনো x থাকবে না। এইভাবে এগোনোর সুবিধা এই যে, gcd বার করতে কোনো root বার করতে হয় না।

> Now $p'(\lambda) = 3\lambda^2 - 4\lambda - 1$
>
> Here $gcd(p(x), p'(x))$ is a constant free of λ.

এই gcd-টা বার করার ধাপগুলো দেখিয়েছি Fig 23-এ। কী করে দুটো polynomial-এর gcd বার করে, ভুলে যাও নি তো? একটা দিয়ে আরেকটাকে ভাগ করে, তারপর remainder-টা (মানে ভাগশেষটা) দিয়ে ফের ভাগ করতে হয়। এইভাবে চলতে থাকে, যতক্ষণ না ভাগ মিলে যায়।

Fig 23

$$\begin{array}{r|l} \text{Divisor: } 3\lambda^2 - 4\lambda - 1 & \text{Quotient: } \tfrac{1}{3}\lambda - \tfrac{2}{9} \\ & \lambda^3 - 2\lambda^2 - \lambda + 6 \\ & -\lambda^3 + \tfrac{4}{3}\lambda^2 + \tfrac{1}{3}\lambda \\ \hline & -\tfrac{2}{3}\lambda^2 - \tfrac{2}{3}\lambda + 6 \\ & \tfrac{2}{3}\lambda^2 - \tfrac{8}{9}\lambda - \tfrac{2}{9} \\ \hline & -\tfrac{14}{9}\lambda + \tfrac{52}{9} \end{array}$$

$$\begin{array}{r|l} \text{Divisor: } -\tfrac{14}{9}\lambda + \tfrac{52}{9} & \text{Quotient: } -\tfrac{27}{14}\lambda - \tfrac{225}{49} \\ & 3\lambda^2 - 4\lambda - 1 \\ & -3\lambda^2 + \tfrac{78}{7}\lambda \\ \hline & \tfrac{50}{7}\lambda - 1 \\ & -\tfrac{50}{7}\lambda + \tfrac{1300}{49} \\ \hline & \tfrac{1251}{49} \end{array}$$

> Thus $p(\lambda)$ and $p'(\lambda)$ have no common root. Hence all the roots of $p(\lambda)$ are distinct.
>
> So all eigenvalues have algebraic multiplicity 1, and hence geometric multiplicity 1.
>
> Thus all the eigenvalues are regular.
>
> So the matrix is diagonalisable.

■

এই অংকটা যদি বুঝে থাকো তবে নীচের অংকটাও করতে পারা উচিত।

Exercise 32: Which of the following is *not* a sufficient condition for a matrix A to be diagonalisable?

(A) All the eigenvalues are simple (*i.e.*, have algebraic multiplicity 1.).

(B) The characteristic polynomial has no repeated root.

(C) $\sum_\lambda |\text{gm}(\lambda, A) - \text{am}(\lambda, A)| = 0$, where the sum is over all distinct eigenvalues of A.

(D) All the eigenvalues of A are either 0 or 1.

■

45.2 Schür decomposition

আমরা জীবন সহজ করার জন্য কোনো square matrix পেলেই তার সঙ্গে similar কোনো সহজ matrix (যেমন diagonal বা triangular) খোঁজার চেষ্টায় ছিলাম। দেখা গেল যে, সব square matrix-ই diagonalisable নয়। ঠিক আছে, আমরা আশা ছাড়ছি না। দেখা যাক কোনো triangular matrix-এর সঙ্গে similarity বার করা যায় কিনা। এইবার উত্তরটা হল--হ্যাঁ, যেকোনো[8] square matrix-ই কোনো না কোনো triangular matrix-এর সঙ্গে similar! মানে তুমি যাই $A_{n\times n}$ দাও না কেন, এমন একটা nonsingular P আর (upper) triangular T খুঁজে পাবে[9], যাতে $A = PTP^{-1}$ হয়। শুধু তাই নয় P-টাকে তুমি চাইলে unitary-ও নিতে পারো, মানে এমন P, যাতে $P^{-1} = P^*$ হয়। একে বলে **Schür decomposition.** এইটা খুবই কাজের কথা, কারণ এর ফলে eigenvalue-ঘটিত অংকের সময়ে তুমি নিশ্চিন্তে প্রথমেই matrix-টাকে triangular ধরে নিতে পারো। ঠিক যেমন fraction নিয়ে কাজ করার সময়ে $\frac{50}{100}$ দেখলে আমরা প্রথমেই ওটাকে $\frac{1}{2}$ করে নিই।

এখানে ছোট্টো করে বলে রাখি যে, এটা একটা অত্যন্ত গুরুত্বপূর্ণ জিনিস হলেও পরীক্ষার সিলেবাসে কিন্তু নেই। সুতরাং নিছক পরীক্ষা-নম্বর-পিপাসুরা এই অংশটা "পরীক্ষার পরে শিখব" বলে মুলতুবি রাখতে চাইতে পারো।

Schur decomposition

Any $A_{n\times n}$ over $\mathbb{C}$ can be written as

$$A = PTP^{-1} = PTP^*$$

for some unitary P and some (upper) triangular T.

[8]এখানে আমরা field নিচ্ছি $\mathbb{C}$.

[9]এখানে upper শব্দটাকে ব্র্যাকেটে রেখেছি, কারণ সেটা গুরুত্বপূর্ণ নয়। একই কথা lower triangular matrix-দের নিয়েও বলা যায়।

এটা আমরা এ বইতে পরে প্রমাণ করব, inner product বলে একটা জিনিস শেখার পরে, নবম অধ্যায়ে গিয়ে। আপাততঃ বলে নিই এর উপকারিতা কী।

45.2.1 উপকারিতা

Schür decomposition দিয়ে বিস্তর জিনিস সহজে প্রমাণ করা যায়। সবগুলো প্রমাণেরই একই কায়দা, তাই "কী করে এগোব?" ভেবে দিশাহারা হতে হয় না। কয়েকটা উদাহরণ দেখলেই ব্যাপারটা বুঝতে পারবে।

Example 36: আমরা শিখেছিলাম যে eigenvalue-দের গুণফল হল determinant-এর সমান। এটা এবার Schür decomposition দিয়ে প্রমাণ করি। প্রথমে লক্ষ্য কর যে, triangular matrix-এর বেলায় এটা বলাই বাহুল্য, কারণ এর diagonal entry-গুলোই হল eigenvalue, আর এই entry-গুলোকে গুণ করলেই determinant হয়। এবার ধর কোনো square matrix নিলাম A, তবে সেটা কোনো একটা triangular matrix T-এর সঙ্গে similar. তাহলে A-র eigenvalue-রা T-এর eigenvalue-র সঙ্গে এক, আবার determinant-ও এক। সুতরাং T-এর ক্ষেত্রে প্রমাণটা হয়েছে মানে A-র ক্ষেত্রেও হয়েছে। সহজ যুক্তি, তাই না? ■

আরেকটা উদাহরণ দেখো।

Example 37: এবার প্রমাণ করব যে trace হল eigenvalue-দের যোগফল। এটাও triangular matrix-দের ক্ষেত্রে বলাই বাহুল্য। এদিকে similarity-র ক্ষেত্রে eigenvalue বা trace কোনোটাই বদলায় না। সুতরাং যেকোনো square matrix-এর ক্ষেত্রেই প্রমাণ হয়ে গেল! ■

আরো একটা প্রমাণ দেখো। এবার একটু কঠিন জিনিস--যদি A যেকোনো square matrix হয় আর $p(x)$ হয় যেকোনো polynomial, তবে $p(A)$-র eigenvalue-রা হবে $p(\lambda_1), ..., p(\lambda_n)$ যেখানে $\lambda_1, ..., \lambda_n$ হল A-র eigenvalue. যেমন A-র eigenvalue-রা যদি $-2, -2, 0, 1, 2$ হয় তবে A^2-এর eigenvalue-রা হবে $4, 4, 0, 1, 4$. লক্ষ্য কর যে, এ থেকে algebraic multiplicity-রও হিল্লে হয়ে যাচ্ছে। যেমন 4 এসেছে 3-বার, দুটো -2 থেকে আর একটা 2 থেকে। সুতরাং A^2-এর জন্য 4-এর algebraic multiplicity হল A-র জন্য -2 আর 2-এর algebraic multiplicity-দের যোগফল। এমনটা যে হয়, সেটা আমরা আগে একটা theorem হিসেবে দিয়েছিলাম, কিন্তু তখনকার বিদ্যে দিয়ে প্রমাণটা করা কঠিন ছিল বলে প্রমাণ করিনি। এবার দেখো Schür decomposition থেকে এটা কেমন সহজে প্রমাণ হয়ে যাবে।

Example 38: ধরো $A_{n\times n}$-র eigenvalue-গুলো হল $\lambda_1, ..., \lambda_n$. (একই eigenvalue একাধিকবার আসতে পারে, algebraic multiplicity যত, ততবার)। যদি $p(x)$ একটা যা খুশি polynomial হয়, তবে দেখাও যে $p(A)$-র eigenvalue-গুলো হবে $p(\lambda_1), ..., p(\lambda_n)$.

SOLUTION: ধাপে ধাপে এগোই--

- প্রথমে ধরো A একটা upper triangular matrix–
  - -- তবে A^2, A^3, ইত্যাদিরাও upper triangular হবে, কারণ upper triangular-দের গুণ করলে upper triangular-ই থাকে। আবার upper triangular-দের যোগ করলেও upper triangular থাকে, সুতরাং $p(A)$-ও upper triangular হবে।
  - -- শুধু তাই নয়, A-র diagonal entry-গুলো যদি a_{ii} হয়, তবে $p(A)$-র diagonal entry-গুলো হবে $p(a_{ii})$. কারণ upper triangular-দের যোগ বা গুণ করলে diagonal entry-রাও একইভাবে যোগ বা গুণ হয়ে যায়।
  - -- Triangular matrix-দের বেলায় diagonal entry-রাই হয় eigenvalue. তাই A-র eigenvalue-রা ছিল a_{ii}-গুলো, আর $p(A)$-র eigenvalue-রা হবে $p(a_{ii})$-গুলো।

সুতরাং upper triangular matrix-দের বেলায় প্রমাণটা হয়ে গেল।

- এবার ধরো A হল যেকোনো square matrix. তাহলে Schür decomposition লাগিয়ে আমরা A-র সঙ্গে similar একটা upper triangular matrix পাব, ধরো T. তাহলে $p(A)$ আর $p(T)$-ও similar হবে। বলা আছে যে A-র eigenvalue-রা $\lambda_1, ..., \lambda_n$. সুতরাং এরাই T-এরও eigenvalue. এক্ষুণি upper triangular-দের জন্য যা প্রমাণ করলাম, তা থেকে পেয়ে যাচ্ছি যে $p(T)$-এর eigenvalue-রা হবে $p(\lambda_1), ..., p(\lambda_n)$. এবং similar হওয়ার সুবাদে $p(A)$-র eigenvalue-ও এরাই। ঠিক সেটাই দেখানোর দরকার ছিল।

■

আমরা আগের একটা অংকে nilpotent বলে একটা ধারণা শিখেছিলাম। ধরো A একটা matrix. একে nilpotent বলব যদি কোনো $k \in \mathbb{N}$-এর জন্য $A^k = O$ হয়। আমরা এটা বলেছিলাম যে, A যদি nilpotent হয়, তবে A-র সব eigenvalue শূন্য হতে বাধ্য। এর উল্টো দিকটাও সত্যিই। সেটা আসে Schür decomposition থেকে--

> THEOREM
>
> A complex matrix is nilpotent if and only if all its eigenvalues are zeroes.

Proof: ধরো matrix-টা হল A. যদি nilpotent হয়, তবে যে eigenvalue-গুলো শূন্য হতে বাধ্য, সেটা বোঝা সহজ--

> Let λ be an eigenvalue of A. Then λ^k is an eigenvalue of $A^k = O$.
>
> So $\lambda^k = 0$. or $\lambda = 0$, as required.

এখানে আমরা এই কথাটা ব্যবহার করেছি যে, শূন্য matrix-এর সব eigenvalue-ই শূন্য হয়।
এবার উল্টোদিকটা। A-র একটা Schür decomposition নিয়ে শুরু করি--

> Let $A = PTP^*$ be a Schür decomposition of A.

তবে A-র eigenvalue-রা হল T-এর diagonal entry-গুলো। যদি তারা সবাই শূন্য হয়, তবে T-কে দেখতে Fig 24-এর মত। এখানে ফাঁকা entry-গুলোতে শূন্য আছে, আর $*$ মানে যা খুশি সংখ্যা।

> Then T is an upper triangular matrix with zeroes in the diagonal.

এইধরণের matrix-দের একটা বিশেষ ধর্ম আছে। যদি T^2, T^3 ইত্যাদি পরপর বার করতে থাকো, তবে দেখতে হবে Fig 25-এর মত। লক্ষ্য করো শূন্যগুলো কেমন কোণাকুণিভাবে সাম্রাজ্য বিস্তার করছে। তাই এইভাবে চলতে চলতে $T^n = O$ হয়ে যেতে বাধ্য।

Fig 24

$$T = \begin{bmatrix} 0 & * & * & * & * & * \\ & 0 & * & * & * & * \\ & & 0 & * & * & * \\ & & & 0 & * & * \\ & & & & 0 & * \\ & & & & & 0 \end{bmatrix}$$

Fig 25

$$T^2 = \begin{bmatrix} 0 & 0 & * & * & * & * \\ & 0 & 0 & * & * & * \\ & & 0 & 0 & * & * \\ & & & 0 & 0 & * \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix} \qquad T^3 = \begin{bmatrix} 0 & 0 & 0 & * & * & * \\ & 0 & 0 & 0 & * & * \\ & & 0 & 0 & 0 & * \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix}$$

> Hence $T^n = O$. So $A^n = PT^nP^* = O$.
> Thus, A is nilpotent, as required.

[Q.E.D]

শূন্যরা যে এইভাবে সাম্রাজ্য বিস্তার করে, সেটা প্রমাণ করা শক্ত নয়। T^2-এর বেলায় করে দেখাচ্ছি। T-এর (i,j)-th entry-কে t_{ij} লিখলে লক্ষ্য করো যে, $i \leq j$ হলে $t_{ij} = 0$ হচ্ছে। এবার T^2-এর (i,j)-th entry হবে

$$\sum_k t_{ik}t_{kj}.$$

এই sum-এর মধ্যে অনেক term-ই শূন্য হবে। যখনই $i \leq k$ তখন $t_{ik} = 0$ আর যখনই $k \leq j$ তখনই $t_{kj} = 0$. এদের বাঁচিয়ে কোনো k সম্ভব একমাত্র তখনই যখন i আর j-র মাঝখানে অন্ততঃ 1 তফাত আছে, যেখানে k-টা ঢুকতে পারে। অর্থাৎ $i \leq j+1$ হলে T^2-এর (i,j)-th entry-টা শূন্য হতে বাধ্য। একইভাবে দেখাতে পারবে যে, T^3-এর বেলায় $i \leq j+2$ লাগবে, T^4-এর বেলায় $i \leq j+3$, ইত্যাদি।

DAY 46
Cayley Hamilton theorem

এবার আমরা একটা theorem-এর কথা শিখব, যার নাম Cayley-Hamilton theorem.

46.1 কী জিনিস?

শুরু করি একটা উদাহরণ দিয়ে--

Example 39: ধরো এই matrix-টা দিলাম, $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. এর characteristic polynomial বার কর, $p(\lambda)$.

এবার দেখো তো $p(A)$ কত হয়!

SOLUTION:

$$p(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 4 \end{vmatrix} = \cdots = \lambda^2 - 5\lambda - 2.$$

সুতরাং

$$p(A) = A^2 - 5A - 2I = \cdots = O.$$

■

এই যে $p(A) = O$ হল, সেটা কিন্তু কাকতালীয় নয়। যেকোনো square matrix-এর বেলাতেই এমনটাই হবে। তুমি যা খুশি A নিয়ে সেটাকে তার নিজের characteristic polynomial-এর পেটের মধ্যে চালান করে দ্যাখো, সব সময়েই O পাবে! এইটাই হল Cayley-Hamilton theorem.

Cayley Hamilton theorem

Let $A_{n\times n}$ be a complex matrix with characteristic polynomial

$$p(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n.$$

Then $p(A) = O$, ie,

$$a_0 I + a_1 A + \cdots + a_n A^n = O.$$

নীচের দুটো অংকে তুমি নিজেই সেটা হাতেকলমে পরীক্ষা করে দেখো।

Exercise 33: Verify Cayley-Hamilton theorem for the matrix A, where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}.$$

[3] **(2020.CC6.3bp1)** ■

Exercise 34: Verify Cayley-Hamilton theorem for the matrix A, where

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 3 & -1 \end{bmatrix}.$$

[3] **(2019.CC6.3ep1)** ■

Example 40: একজন ছাত্র এই theorem-টাকে এইভাবে প্রমাণ করার চেষ্টা করেছে--সে জানে $p(\lambda) = |\lambda I - A|$. সুতরাং যদি দুইদিকেই λ-র জায়গায় A বসিয়ে দিই, তবে বাঁদিকটা হয়ে যাবে $p(A)$, আর ডানদিকটা হবে $|A - AI| = |O| = 0$. ব্যস্, সে ভারী খুশি হয়ে ঘোষণা করে দিল যে প্রমাণটা হয়ে গেছে! ও কি ঠিক করেছে?

SOLUTION: না, ও যেটা "প্রমাণ" করেছে, সেটার মানেই দাঁড়ায় না। ধরো A হল 2×2. তাহলে ওর দাবী হল $p(A)$ হচ্ছে 0. কিন্তু $p(A)$-র সাইজ হল 2×2, আর 0-টা এসেছে একটা determinant থেকে, সুতরাং ওটা খালি 0 সংখ্যা, O matrix নয়! একটা 2×2 সাইজের matrix কী করে একটা সংখ্যার সমান হয়?

আসলে এখানে $p(\lambda) = |\lambda I - A|$ কথাটা খালি scalar কোনো λ-র ক্ষেত্রেই খাটে। ওখানে λ-র জায়গায় কোনো matrix বসানো চলে না, কারণ সেক্ষেত্রে বাঁদিকটা matrix হয়ে যায়, যদিও ডানদিকটা একটা সংখ্যাই থেকে যায়! ■

যাই হোক, এই অতি সহজ "প্রমাণ"-টা কাজ করল না বটে, কিন্তু তা বলে প্রমাণটা খুব কিছু কঠিন নয়। প্রমাণটা শুরু হয় adjugate ব্যবহার করে। কোনো matrix-এর adjugate কী জিনিস মনে আছে তো? Matrix-টার প্রত্যেকটা entry-র cofactor বার করে সেই cofactor-দের matrix-টার transpose নিলে adjugate পাওয়া যায়। এদের নিয়ে সবচেয়ে গুরুত্বপূর্ণ জিনিস যেটা শিখেছিলাম সেটা হল--

> <u>Proof:</u> We know that for any square matrix B
>
> $$B \cdot \text{adj}(B) = |B| \cdot I.$$

এটা হবার জন্য B যেকোনো square matrix হলেই চলে। এবার আমরা B-এর জায়গায় $\lambda I - A$ বসাব--

> For $B = \lambda I - A$,
>
> $$(\lambda I - A)\text{adj}(\lambda I - A) = |\lambda I - A| \cdot I = p(\lambda) I. \quad (*)$$

এখানে একটা উদাহরণ নিলে সুবিধা হবে, ধরো

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & -1 & 4 \end{bmatrix}.$$

তাহলে

$$\lambda I - A = \begin{bmatrix} \lambda - 1 & -2 & -3 \\ -4 & \lambda - 5 & -6 \\ -3 & 1 & \lambda - 4 \end{bmatrix}.$$

লক্ষ্য কর যে, এর diagonal entry-গুলো "$\lambda-$কিছু একটা" চেহারার, আর বাকি entry-গুলো কিছু সংখ্যা। এর adjugate-এর entry-গুলো তাহলে কীরকম দেখতে হবে? ওরা প্রত্যেকে একেকটা cofactor, মানে একেকটা 2×2 সাইজের submatrix-এর determinant (সঙ্গে একটা $+$ বা $-$ চিহ্ন লাগানো)। যেমন, এই adjugate-এর $(1,1)$-th entry-টা হবে

$$(\lambda - 5)(\lambda - 4) + 6.$$

যদি A-টা $n \times n$ সাইজের হয়, তবে একইভাবে প্রতিটা cofactor-ই হবে λ-র একেকটা polynomial, এবং কারো degree-ই $(n-1)$-এর বেশী হতে পারে না।

> Now each entry of $\text{adj}(\lambda I - A)$ is a polynomial of degree at most $(n-1)$ in λ.
> Collecting powers of λ we can write $\text{adj}(\lambda I - A)$ as a polynomial of degree $n-1$ in λ with matrix coefficients:
>
> $$\text{adj}(A - \lambda I) = H_0 + H_1\lambda + \cdots + H_{n-1}\lambda^{n-1}.$$

এই ধাপটায় এসে অনেকে খাবি খায়। কিন্তু এটা একটু সাজিয়ে লেখা ছাড়া আর কিছুই নয়। একটা 2×2 উদাহরণ নিয়ে দেখাই। ধরো $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$. এর জন্য

$$\text{adj}(\lambda I - A) = \begin{bmatrix} \lambda - 2 & 3 \\ 2 & \lambda - 1 \end{bmatrix}.$$

এর মধ্যে λ ছাড়া অংশ, আর λ-ওয়ালা অংশটাকে আলাদা করে লিখব--

$$\text{adj}(\lambda I - A) = \underbrace{\begin{bmatrix} -2 & 3 \\ 2 & -1 \end{bmatrix}}_{H_0} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{H_1}\lambda.$$

এবার আমরা (*)-এর দিকে চোখ ফেরানো যাক। ওর দুপাশেই λ-র দুটো polynomial আছে, যাদের coefficient-গুলো হল একেকটা matrix. ডানদিকে আছে--

> The RHS of (*) is
>
> $$p(\lambda)I = p_0 I + p_1 I\lambda + \cdots + p_n I\lambda^n,$$

আর বাঁদিকে আছে--

> The LHS is
>
> $$\begin{aligned}(\lambda I - A)\text{adj}(A-\lambda I) &= (\lambda I - A)(H_0 + H_1\lambda + \cdots + H_{n-1}\lambda^{n-1}) \\ &= -AH_0 + (H_0 - AH_1)\lambda + (H_1 - AH_2)\lambda^2 + \\ &\qquad \cdots + (H_{n-2} - AH_{n-1})\lambda^{n-1} + H_{n-1}\lambda^n.\end{aligned}$$

যেহেতু এই দুই পাশ সমান, তাই--

> So equating coefficients of powers of λ in (*),
>
> $$\begin{aligned}p_0 I &= -AH_0 \\ p_1 I &= H_0 - AH_1 \\ p_2 I &= H_1 - AH_2 \\ &\vdots \\ p_{n-1} I &= H_{n-2} - AH_{n-1} \\ p_n I &= H_{n-1}\end{aligned}$$

এবার এদেরকে বাঁদিক থেকে $A, A^2, A^3, \ldots$ ইত্যাদি দিয়ে গুণ করে দেব। খালি প্রথম equation-টাকে কিছু করব না, বা ভাবতে পারো যেন ওটাকে $A^0 = I$ দিয়ে গুণ করছি।

> Thus, multiplying by A from left,
>
> $$\begin{aligned}p_0 I &= -AH_0 \\ p_1 A &= AH_0 - A^2H_1 \\ p_2 A^2 &= A^2H_1 - A^3H_2 \\ &\vdots \\ p_{n-1}A^{n-1} &= A^{n-1}H_{n-2} - A^nH_{n-1} \\ p_nA^n &= A^nH_{n-1}\end{aligned}$$

এবার পরপর যোগ করে দিলেই দেখবে ঘ্যাঁচাঘ্যাঁচ্ সব কেটে যাচ্ছে!

> Adding, $p(A) = O$, as required. [QED]

46.2 প্রয়োগ

Cayley-Hamilton theorem-টা দেখতে ভালোই, তবে খুব যে একটা প্রয়োগ আছে এমন নয়। আমরা এখানে দুটো জিনিস আলোচনা করব যেগুলোকে খানিকটা কষ্ট-কল্পনা করে প্রয়োগ বলে ভাবা যায়।

46.2.1 বড় power-দের ছোটো power দিয়ে লেখা

ধরো একটা 3×3 সাইজের matrix আছে, যার characteristic polynomial হল--

$$p(\lambda) = \lambda^3 - 2\lambda^2 + 3\lambda + 10.$$

তাহলে Cayley-Hamilton theorem থেকে পাব $p(A) = A^3 - 2A^2 + 3A + 10I = O$, মানে--

$$A^3 = 2A^2 - 3A - 10I.$$

অতএব A^3-কে I, A, A^2-এর একটা linear combination হিসেবে লিখে ফেলা গেল। এবার দুই পাশকে A দিয়ে গুণ করলে A^4-কে লিখে ফেলা যাবে A, A^2, A^3-এর linear combination হিসেবে। যেহেতু A^3-কে I, A, A^2 দিয়ে লেখা যাচ্ছিলই, সুতরাং সেটা বসিয়ে দিলে A^4-কেও I, A, A^2-এর linear combination হিসেবে লেখা যাবে। একইভাবে $A^5, A^6, \ldots$ ইত্যাদি সকলকেই খালি I, A, A^2 দিয়ে লেখা যাবে। ছোটোবেলায় লুডো খেলার সময়ে একটা নিয়ম ছিল--তিন ছক্কা পড়লে কেটে যায়। ফলে যতই ছক্কা পড়ুক না কেন, তার প্রভাবটা দুই ছক্কার উপরে কিছুতেই উঠত না, এও কতকটা সেইরকম।
এইভাবে এগোলে A^n-কে কেবল I, A, A^2-এর linear combination হিসেবে লেখা যাবে বটে, কিন্তু সেই linear combination-এ ব্যবহৃত coefficient-গুলো বার করা খুব একটা সহজ নাও হতে পারে। এবার একটা অংক দেখব, যেখানে এই ধারণাটার প্রয়োগ রয়েছে।

Example 41: If

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

show that for every integer $n(\geq 3)$

$$A^n = A^{n-2} + A^2 - I.$$

Hence evaluate A^{50}.
SOLUTION: যদি n-ওয়ালা জিনিসগুলোকে একদিকে নিয়ে যাই, তবে আমাদের দেখাতে হবে $A^n - A^{n-2} = A^2 - I$. এই চেহারাটা বেশ সুন্দর, বাঁদিকে A-র power-এর পার্থক্য $n - (n-2) = 2$, আবার ডানদিকেও তাই, $2 - 0 = 2$. এই প্যাটার্ণটা মনে রেখো, শীঘ্রই কাজে লাগবে।

> A has characteristic polynomial
>
> $$p(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ -1 & \lambda & -1 \\ 0 & -1 & \lambda \end{vmatrix} = \cdots = \lambda^3 - \lambda^2 - \lambda + 1.$$
>
> By Cayley Hamilton theorem, $p(A) = O$, ie,
>
> $$A^3 - A^2 - A + I = O,$$
>
> or
>
> $$A^3 - A = A^2 - I.$$

সেই প্যাটার্ণটা দেখতে পাচ্ছ? বস্তুতঃ, যেটা দেখাতে বলেছিল তার $n = 3$ কেসটা এটাই। এবার n বাড়ানো যাক--

> Now
>
> $$\begin{aligned} A^4 - A^2 &= A(A^3 - A) = A(A^2 - I) = A^3 - A = A^2 - I \\ A^5 - A^3 &= A(A^4 - A^2) = A(A^2 - I) = A^3 - A = A^2 - I \\ &\vdots \end{aligned}$$

Continuing like this, we get $A^n - A^{n-2} = A^2 - I$, or

$$A^n = A^{n-2} + A^2 - I,$$

as required.

এবার দ্বিতীয় অংশ--

For $n = 50$,

$$\begin{aligned} A^{50} &= A^{48} + A^2 - I = A^{46} + 2(A^2 - I) \\ &= \cdots \\ &= A^2 + 24(A^2 - I) \\ &= 25A^2 - 24I \\ &= 25\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} - 24\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 25 & 1 & 0 \\ 25 & 0 & 1 \end{bmatrix}. \end{aligned}$$

■

এইবার একটু কঠিন একটা অংক করি।

Exercise 35: Let A be any $n \times n$ matrix. Consider the set

$$\{p(A) \ : \ p \text{ any polynomial}\}.$$

Show that this is a vector space of dimension $\leq n$. ■

46.2.2 Inverse বার করা

Cayley-Hamilton theorem-এর আরেকটা প্রয়োগ হল কোনো matrix-এর inverse বার করা। কায়দাটা একটা উদাহরণ দিয়ে দেখিয়েছি নীচের অংকটায়। আগেই বলে রাখি যে, inverse বার করার পক্ষে এটা মোটেই খুব সহজ কায়দা নয়, নীচের অংকে matrix-টা নেহাত ছোটো বলে অসুবিধা হবে না। কিন্তু বড় বড় matrix-এর বেলায় Gauss-Jordan elimination লাগিয়ে inverse বার করা অনেক বেশী সহজ।

Example 42: Use Cayley Hamilton theorem to compute A^{-1} where

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}.$$

SOLUTION:

The characteristic polynomial of A is

$$p(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 2 & -1 \\ -3 & \lambda - 5 \end{vmatrix} = \lambda^2 - 7\lambda + 7.$$

So by Cayley Hamilton theorem,

$$A^2 - 7A + 7I = O,$$

এবার constant term-টা, মানে যে term-টায় I আছে, সেটাকে একদিকে নিয়ে আসব--

or $A^2 - 7A = -7I$, or

$$A(A - 7I) = (A - 7I)A = -7I,$$

আমরা ডানদিকে খালি I রাখতে চাই, তাই -7 দিয়ে ভাগ করে দিই--

or

$$A\left(-\tfrac{1}{7}(A - 7I)\right) = \left(-\tfrac{1}{7}(A - 7I)\right)A = I.$$

তার মানে $-\frac{1}{7}(A - 7I)$ এমন একটা matrix, যা দিয়ে A-কে বাঁদিক বা ডানদিক থেকে গুণ করে দিলে I পাওয়া যাচ্ছে। তার মানেই তো হল--

So

$$A^{-1} = -\tfrac{1}{7}(A - 7I) = \cdots = \tfrac{1}{7}\begin{bmatrix} 5 & -1 \\ -3 & 2 \end{bmatrix}.$$

■

এই কায়দাটা কি সব সময়েই খাটবে? আমরা যেটা করলাম, সেটা হল I বাদে বাকি সবগুলো term-কে বাঁদিকে নিয়ে গেলাম। যদি A-টা $n \times n$ হয়, তবে এরা হল $A, A^2, ..., A^n$-ঘটিত term-গুলো। এদের সবার মধ্যেই A আছে, সেটাকে "কমন" নিয়ে লিখতে পারলাম

$$A \times \text{কিছু একটা} = I,$$

এবং তা থেকেই ওই "কিছু একটা"-টাকে A^{-1} বলে দাবি করতে পারলাম। প্রশ্ন হল--এমনটা কি সবসময়েই করা যাবে? যদি মনে করো characteristic polynomial-টায় কোনো constant term না থাকত? তবে তো কোনো I পেতামই না। সে ক্ষেত্রে কী হত?

নীচের অংকে অভয় দিয়ে বলা হয়েছে যে, যেকোনো nonsingular matrix-এর জন্যই কায়দাটা কাজ করবে।

Example 43: Write, with proof, how the inverse can be found for any nonsingular matrix using Cayley-Hamilton theorem.

SOLUTION:

Let A be a nonsingular matrix. Then $|A| \neq 0$.

$\therefore 0$ is not an eigenvalue of A.

Thus, if

$$p(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n,$$

> is the characteristic polynomial of A, then $a_0 = p(0) \neq 0$.

সুতরাং constant term-টা শূন্য হয়ে যাওয়ার ভয় রইল না। যদি কখনো দেখো $p_0 = 0$ হয়ে গেছে, অমনি জানবে A-টা nonsingular-ই নয়।

> By Cayley Hamilton theorem,
>
> $$a_0 I + a_1 A + \cdots + a_n A^n = O,$$
>
> or, since $a_0 \neq 0$,
>
> $$A\left[-\tfrac{1}{a_0}(a_1 I + a_2 A + \cdots + a_n A^{n-1})\right] = \left[-\tfrac{1}{a_0}(a_1 I + a_2 A + \cdots + a_n A^{n-1})\right] A = I.$$
>
> Hence
>
> $$A^{-1} = -\tfrac{1}{a_0}(a_1 I + a_2 A + \cdots + a_n A^{n-1}).$$

■

লক্ষ্য করো যে, এই কায়দায় শুধু যে A^{-1} বার করা গেল তাই নয়, সেটাকে আবার $I, A, A^2, ..., A^{n-1}$-এর একটা linear combination হিসেবে লেখাও গেল।

Exercise 36: Verify Cayley-Hamilton theorem for the matrix A. Also express A^{-1} as a polynomial in A and then compute A^{-1} where

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 3 & -1 \end{bmatrix}.$$

[3+1+1] **(2019.CC6.3efull)**
HINT: প্রথম অংশটা আজকে একটু আগেই করতে দিয়েছিলাম। সেখানে characteristic polynomial-ও বার করেছিলে। এবার সেখান থেকে A^{-1} বার করতে আর কতক্ষণ?
■

Exercise 37: Verify Cayley-Hamilton theorem for the matrix A. Also express A^{-1} as a polynomial in A and then compute A^{-1} where

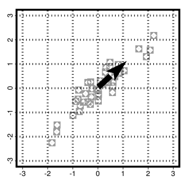
$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}.$$

[5] **(2020.CC6.3b)**
HINT: এটারও প্রথম অংশটা তুমি আজকে একটু আগেই দেখেছ।
■

Exercise 38: Use Cayley-Hamilton theorem to compute A^{-1}, where $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$.[3]
(2022.CC6.3fi) ■

Answers

1. $(\lambda-1)(\lambda-3)(\lambda-6)$. **2.** Eigenvalue-গুলো হল $2, 3$ এবং 5. এদের জন্য eigenvector-রা হল যথাক্রমে $(x, 0, 0)$, $(0, x, 0)$ এবং $(0, 0, x)$ জাতীয় যেকোনো vector, যেখানে $x \neq 0$. **3.** Eigenvalue-রা হল $1, 1$ এবং -1. দুবার 1 লিখলাম কারণ এখানে characteristic polynomial-টা হল $(\lambda-1)^2(\lambda+1)$. লক্ষ্য করো, $(\lambda-1)$ রয়েছে দুবার। 1-এর জন্য eigenvector-রা সবাই (a, a, b)-র মত দেখতে nonzero vector. আর -1-এর জন্য eigenvector-রা দেখতে $(a, -a, 0)$-র মত দেখতে nonzero vector-রা। Eigenspace-গুলো হল যথাক্রমে span$\{(1,1,0), (0,0,1)\}$ এবং span$\{(1,-1,0)\}$. **4.** (D). **5.** $\begin{bmatrix} 0 & \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$. **6.** সব দিকে সমান কাপের মত, যার প্রস্থচ্ছেদ হবে নিখুঁত গোল। **7.** Eigenvalue দুটো আসছে 1.387 আর 0.056. প্রথমটার জন্য একটা unit eigenvector হল $(0.722, 0.692)$. তীর চিহ্ন আঁকলে দেখাবে এরকম--



8. Eigenvalue-গুলো হল -1 আর 1. এদের algebraic multiplicity যথাক্রমে 2 এবং 1. এদের দুজনেরই geometric multiplicity হল 1. **9.** ঠিক, কারণ $1 \leq \mathrm{gm}(\lambda, A) \leq \mathrm{am}(\lambda, A) = 1$. তাই $\mathrm{gm}(\lambda, A) = \mathrm{am}(\lambda, A)$. **10.** অসম্ভব। একটা diagonal matrix-এর বেলায় যে কোনো eigenvalue-এর জন্যই algebraic multiplicity আর geometric multiplicity সমান হয়। **11.** ধরো eigenvalue-টাকে λ নাম দিলাম। তাহলে $A(\vec{e}_1 + \vec{e}_2) = \lambda(\vec{e}_1 + \vec{e}_2)$. কিন্তু $A(\vec{e}_1 + \vec{e}_2) = a_1\vec{e}_1 + a_2\vec{e}_2$. তাই $\lambda(\vec{e}_1 + \vec{e}_2) = a_1\vec{e}_1 + a_2\vec{e}_2$. সুতরাং প্রথম entry থেকে পাই $\lambda = a_1$,আর দ্বিতীয় entry থেকে $\lambda = a_2$. সুতরাং $a_1 = a_2$. **12.** $\lambda_1 - \lambda_2 = 0$, কারণ $A(\vec{e}_1 + \vec{e}_2) = a_1\vec{e}_1 + a_2\vec{e}_2 = \lambda_1(\vec{e}_1 + \vec{e}_2)$. সুতরাং $a_1 = a_2 = \lambda_1$ হবেই, যেহেতু $\{\vec{e}_1, \vec{e}_2\}$ হল linearly independent. একইভাবে পরের শর্ত থেকে পাই $a_2 = a_3 = \lambda_2$. সুতরাং $\lambda_1 = a_2 = \lambda_2$. **13.** (A), কারণ $n-k$ হল শূন্য eigenvalue-র algebraic multiplicity, আর $n-r$, মানে nullity(A) হল তার geometric multiplicity. এদিকে আমরা জানি যে algebraic multiplicity$\geq$ geometric multiplicity হয়। সর্বদাই যে $k = r$ নাও হতে পারে, তার একটা counterexample হল $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. এর rank হল 1, কিন্তু nonzero eigenvalue একটাও নেই। **14.**

> Let λ be an eigenvalue of Hermitian A with eigenvector $\vec{v} \neq \vec{0}$.
>
> Then $\vec{v}^*A\vec{v} = \lambda\vec{v}^*\vec{v} = \lambda\|\vec{v}\|^2$.
>
> Also $\vec{v}^*A\vec{v} = (A^*\vec{v})^*\vec{v} = (A\vec{v})^*\vec{v} = \overline{\lambda}\vec{v}^*\vec{v} = \overline{\lambda}\|\vec{v}\|^2$.
>
> Hence $\lambda\|\vec{v}\|^2 = \overline{\lambda}\|\vec{v}\|^2$.
>
> So, since $\|\vec{v}\| \neq 0$, we have $\lambda = \overline{\lambda}$.
>
> If $\lambda = a + ib$ for $a, b \in \mathbb{R}$, then $a + ib = a - ib$. So $b = 0$.
>
> Hence λ is real.

15.

> Let λ be an eigenvalue of skew-Hermitian A with eigenvector $\vec{v} \neq \vec{0}$.
> Then $\vec{v}^* A\vec{v} = \lambda\vec{v}^*\vec{v} = \lambda\|\vec{v}\|^2$.
> Also $\vec{v}^* A\vec{v} = (A^*\vec{v})^*\vec{v} = -(A\vec{v})^*\vec{v} = -\overline{\lambda}\vec{v}^*\vec{v} = -\overline{\lambda}\|\vec{v}\|^2$.
> Hence $\lambda\|\vec{v}\|^2 = -\overline{\lambda}\|\vec{v}\|^2$.
> So, since $\|\vec{v}\| \neq 0$, we have $\lambda = -\overline{\lambda}$.
> If $\lambda = a+ib$ for $a, b \in \mathbb{R}$, then $a+ib = -a+ib$. So $a = 0$.
> Hence λ is purely imaginary or zero.

16. ধরো eigenvalue তিনটে হল $a, b, 1$ ($a = b$ হতে পারে)। তবে determinant হল $ab = 6$. আর trace হল $a+b+1 = 6$. এ দুটো equation মিলিয়ে আসছে $a^2 - 5a + 6 = 0$, বা $a = 2$ বা $a = 3$. যদি $a = 2$ হয়, তবে $b = 3$, আর যদি $a = 3$ হয়, তবে $b = 2$. সুতরাং eigenvalue দুটো 2 আর 3. **17.**

> Let $\lambda_1, ..., \lambda_k$ be the distinct eigenvalues with corresponding eigenspaces $E_1, ..., E_k$.
> Then $\dim(E_i) = \mathrm{gm}(\lambda_i, A) = \mathrm{am}(\lambda_i, A)$, since the eigenvalues are regular.
> Hence $\sum \dim(E_i) = \sum \mathrm{am}(\lambda_i, A) = n$.
> We know that $\forall \neq j \quad E_i \cap E_j = \{\vec{0}\}$. Let $\{\vec{v}_1^t, ..., \vec{v}_{g_t}^t\}$ be any basis of E_t. Then their union $\{\vec{v}_{g_t}^t \ : \ t = 1, ..., k\}$ has size n.
> Shall show that this set is linearly independent.
> Let $\sum_t \sum_i a_i^t \vec{v}_i^t = \vec{0}$.
> For $t = 1, ..., k$, let $\vec{u}_t = \sum_i a_i^t \vec{v}_i^t$.
> Then $\sum_t \vec{u}_t = \vec{0}$.
> Hence $\vec{u}_1, ..., \vec{u}_k$ must all be $\vec{0}$.
>
> ⟦ Because:
> The nonzero $\vec{u}_t$'s are eigenvectors corresponding to the distinct eigenvalues λ_t's., and hence are linearly independent. So their sum cannot be $\vec{0}$.
> ⟧
>
> Since $\vec{u}_t = \vec{0}$, hence $\sum_i a_i^t \vec{v}_i^t = \vec{0}$, and hence $\forall i \quad a_i^t = 0$, as $\{\vec{v}_i^t \ : \ i = 1, ..., g_t\}$ is linearly independent.
> So $\{\vec{v}_{g_t}^t \ : \ t = 1, ..., k\}$ is linearly independent. Since its size equals n, which is the dimension of $\mathbb{C}^n$, this must be a basis.

18.

> Take any $\vec{v} \in \mathscr{C}(p(A))$. Then $\vec{x} = p(A)\vec{z}$ for some $\vec{z}$. So $A\vec{x} = Ap(A)\vec{z}$.
> Since $Ap(A) = p(A)A$, we have $A\vec{x} = p(A)(A\vec{z}) \in \mathscr{C}(p(A))$, as required.

19.

> Let $\vec{x} \in \mathscr{N}(p(A))$. Then $p(A)\vec{x} = \vec{0}$.
>
> Now $p(A)(A\vec{x}) = Ap(A)\vec{x} = 0$. So $A\vec{x} \in \mathscr{N}(p(A))$, as required.
>
> Counter-example for $\mathscr{R}(p(A))$: Take $p(x) = x$ and $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $\vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathscr{R}(p(A))$.
>
> But $A\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \notin \mathscr{R}(p(A))$.

20. $A\vec{\mathbb{1}} = \alpha\vec{\mathbb{1}}$, তাই α একটা eigenvalue আর তার জন্য $\vec{\mathbb{1}}$ একটা eigenvector. যেহেতু B'-এর row sum-গুলো সব α, তাই B'-এর একটা eigenvalue হবে α. সুতরাং B-এরও একটা eigenvalue হবে α. লক্ষ্য কর যে, B-এর ক্ষেত্রে কিন্তু eigenvector নিয়ে কোনো দাবী করছি না।
21. $(AB)\vec{v} = A(B\vec{v}) = A(\mu\vec{v}) = \mu A\vec{v} = \lambda\mu\vec{v}$. **22.** হ্যাঁ হবে, algebraic multiplicity হবে $1 + (7-5) = 3$. এর কারণ হল $p_{BA}(\lambda) = \lambda^{7-5}p_{AB}(\lambda)$. এদিকে $p_{AB}(\lambda)$-র মধ্যে তো λ একটা factor ছিলই, তার সাথে এখন আবার $\lambda^{7-5} = \lambda^2$ এসে লেগেছে। **23.** $A^n\vec{v} = A(A(\cdots(A\vec{v})\cdots)) = \underbrace{\lambda\cdot\lambda\cdots\lambda}_{n \text{ times}}\vec{v} = \lambda^n\vec{v}$.
24. $(\lambda^2 I - A^2)\vec{v} = \vec{0}$. সুতরাং $(\lambda I + A)(\lambda I - A)\vec{v} = \vec{0}$. কিন্তু $-\lambda$ যেহেতু A-র একটা eigenvalue নয়, তাই $(\lambda I + A)$ একটা nonsingular matrix. তাই $(\lambda I - A)\vec{v} = (\lambda I + A)^{-1}\vec{0} = \vec{0}$. অর্থাৎ $A\vec{v} = \lambda\vec{v}$.
25. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$-এর একটাই eigenvalue, সেটা হল $\lambda_1 = 0$. এদিকে $A^2 = O$. সুতরাং $\mu = 0$ এবং $\text{gm}(0, A^2) = 2$. কিন্তু $\sum_{i=1}^t \text{gm}(\lambda_i, A) = \text{gm}(0, A) = 1$. **26.** $A^2 - A = O$, তাই λ যদি A-র একটা eigenvalue হয় তবে $\lambda^2 - \lambda = 0$, তাই $\lambda = 0$ বা 1. **27.** সব eigenvalue-ই 0 হতে বাধ্য। **28.**

> Since A is skew symmetric, all its eigenvalues are either 0 or purely imaginary, ie, of the form $i\lambda$, for $\lambda \in \mathbb{R}$.
>
> Also the eigenvalues of A^2 are the squares of the eigenvalues of A, ie, of the form $(i\lambda)^2 = -\lambda^2$.
>
> Thus A^2 cannot have imaginary characteristic roots.

29. $|C + \lambda I| = |BAB^{-1} + \lambda I| = |B(A + \lambda I)B^{-1}| = |B| \cdot |A + \lambda I| \cdot |B^{-1}| = |A + \lambda I|$. **30.** এখানে $P\vec{v}_i$-রা linearly independent হবে, কারণ যদি $\sum c_i P\vec{v}_i = \vec{0}$ হয়, তবে $P\sum c_i\vec{v}_i = \vec{0}$, এবং P যেহেতু nonsingular, তাই $\sum c_i\vec{v}_i = \vec{0}$. সুতরাং c_i-র সবাই 0, যেহেতু $\vec{v}_i$-রা linearly independent. যদি $B\vec{w} = \lambda\vec{w}$ হয়, তবে $PAP^{-1}\vec{w} = \lambda\vec{w}$, মানে $AP^{-1}\vec{w} = \lambda P^{-1}\vec{w}$. আবার $\vec{w} \neq \vec{0}$ এবং P-টা nonsingular হওয়ায় $P^{-1}\vec{w} \neq \vec{0}$. সুতরাং $P^{-1}\vec{w}$ হল A-র একটা eigenvector, যেখানে eigenvalue-টা λ. সুতরাং $P^{-1}\vec{w}$-কে $P^{-1}\vec{w} = \sum d_i\vec{v}_i$ আকারে লেখা যাবে। অতএব $\vec{w} = \sum d_i P\vec{v}_i$. সুতরাং generating set দেখানো হয়ে গেল।
31. এখানে দুটো eigenvalue আছে, 4 আর 1. এদের algebraic multiplicity যথাক্রমে 1 এবং 2. প্রথমটার জন্য একটা eigenvector হল $(1, -1, 1)$, এবং দ্বিতীয়টার eigenspace-এর একটা basis হল $\{(1, 0, -1), (0, 1, 1)\}$. সুতরাং যদি $P = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ নিই, তবে $P^{-1}AP = \text{diag}(4, 1, 1)$ হবে। এছাড়াও আরো অন্য বহু উত্তর সম্ভব।
32. (D). উদাহরণস্বরূপ $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$-র দুটো eigenvalue-ই 0, কিন্তু এটা diagonalisable নয়। **33.** এখানে

characteristic polynomial-টা হল $p(x) = x^3 - 4x^2 + 2x + 1$. সরাসরি কষলেই আসছে $p(A) = O$.
34. এখানে characteristic polynomial-টা হল $p(x) = x^3 + x^2 - 8x - 9$. সরাসরি কষলেই আসছে $p(A) = O$.
35. যাবতীয় $n \times n$ সাইজের matrix-রা মিলে তো একটা vector space হয়ই। যেহেতু দুটো polynomial-কে যোগ-বিয়োগ করলে বা কোনো সংখ্যা দিয়ে গুণ করলে আবার একটা polynomial হয়, তাই V হবে এই vector space-টার একটা subspace. যেহেতু যেকোনো A^k-কেই $I, A, ..., A^{n-1}$-এর linear combination হিসেবে লিখে ফেলা যায়, তাই $\{I, A, ..., A^n\}$ হল এই subspace-এর একটা generating set. এর সাইজ যখন n, সুতরাং dim অবশ্যই $\leq n$ হতে বাধ্য। **36.** এখানে characteristic polynomial হল $x^3 + x^2 - 8x - 9$. তাই $A^{-1} = \frac{1}{9}(A^2 + A - 8I) = \cdots = \begin{bmatrix} -\frac{2}{9} & \frac{5}{9} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & 0 \\ \frac{5}{9} & \frac{1}{9} & -\frac{1}{3} \end{bmatrix}$. **37.** এখানে characteristic polynomial হল $x^3 - 4x^2 + 2x + 1$. তাই $A^{-1} = -(A^2 - 4A + 2I) = \cdots = \begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 1 \\ 1 & 3 & -2 \end{bmatrix}$. **38.** এখানে characteristic polynomial হল $p(x) = x^3 + x^2 - 3x - 3$. তাই $A^{-1} = \frac{1}{3}(A^2 + A - 3I) = \cdots = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3}. \end{bmatrix}$

Chapter VIII
Inner product space

DAY 47
Inner product (part 1)

আমরা এই বইতে ইতিপূর্বে বারবারই দেখেছি যে basis মানে আসলে coordinate system. উদাহরণস্বরূপ $\mathbb{R}^2$-এর দুটো basis দেখিয়েছি Fig 1(a) আর Fig 1 (b)-তে। এদেরকে দিয়ে তৈরী coordinate system ধরে যে দুরকমের গ্রাফ কাগজ হয়, তাদের ছবি রয়েছে Fig 2 (a) আর Fig 2 (b)-তে। এই দুই ধরণের গ্রাফ কাগজের মধ্যে কোনটা নিয়ে কাজ করা বেশী সুবিধাজনক? নিশ্চয়ই দ্বিতীয়টা! কারণ আমরা 90° কোণ নিয়ে কাজ করতেই বেশী স্বচ্ছন্দ বোধ করি। সেই কারণেই আমাদের বাড়িঘরের দেওয়াল থাকে মেঝের সঙ্গে 90° কোণে। এই সুবিধার জন্যই ছোটোবেলায় স্কুলে যেসব গ্রাফ দেখেছ, সেগুলোতে axis-গুলো পরস্পরের সঙ্গে 90° কোণ করে নেওয়া হত। একই কারণে আমরাও সেই রকম basis নিয়ে কাজ করতে ভালোবাসি, যেখানে vector-গুলো পরস্পরের সঙ্গে 90° কোণে থাকে। যদি $\mathbb{R}^2$ বা $\mathbb{R}^3$ নিয়ে কাজ করি, তবে দুটো vector-এর মধ্যে কোণ কাকে বলে সেটা চোখে দেখেই বোঝা যায়। কিন্তু যদি অন্য vector space নিয়ে কাজ করি, যেমন $\mathbb{R}^4$ বা $\mathbb{R}^{100}$? বা polynomial-দের দিয়ে তৈরী কোনো vector space? সেখানেও কি কোণের ধারণাটা সম্ভব? উত্তর হল--হ্যাঁ, এবং সেই কাজে প্রধান সহায় হল **inner product**, যেটা আমরা এখন শিখব।

ধরো তোমাকে $\mathbb{R}^2$-র মধ্যে দুটো vector দিলাম $\vec{a}$ আর $\vec{b}$. দিয়ে জানতে চাইলাম এরা পরস্পরের সঙ্গে perpendicular (লম্ব) কিনা। Linear algebra-য় অবশ্য আমরা "perpendicular" শব্দটা ব্যবহার না করে "**orthogonal**" শব্দটা ব্যবহার করি[1]। ওরা orthogonal কিনা সেটা পরীক্ষা করার একটা কায়দা হল Pythagoras' theorem ব্যবহার করা। এর জন্য প্রথমে vector দুটোকে তীর চিহ্ন দিয়ে এঁকে তাদের মাথা দুটোকে আরেকটা তীর চিহ্ন দিয়ে জুড়ে দাও। তাহলে একটা triangle পেলে। Fig 3 দ্যাখো। এর বাহুগুলো হল তিনটে vector. দুটো তো জানিই, $\vec{a}$ আর $\vec{b}$. তৃতীয়টা হল $\vec{c} = \vec{a} - \vec{b}$. এদের দৈর্ঘ্য হল যথাক্রমে

$$\sqrt{a_1^2 + a_2^2}, \quad \sqrt{b_1^2 + b_2^2} \text{ এবং } \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

[1]আমরা orthogonal matrix-দের কথা শুনেছি। এখানে orthogonal vector-দের কথা শিখব।

Fig 1

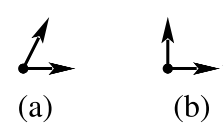


Fig 2

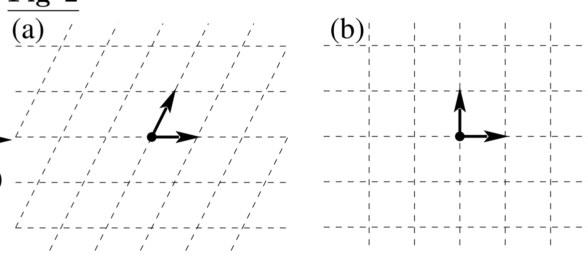
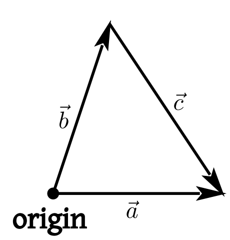


Fig 3



এই ফর্মুলাগুলো দেখতে বড্ড খটমট, এদের একটা ভদ্র নাম দিই, $\vec{a}$-এর দৈর্ঘ্যকে বলব **norm** of $\vec{a}$, আর লিখব $\|\vec{a}\|$, একইভাবে বাকি দুটো দৈর্ঘ্য হবে $\|\vec{b}\|, \|\vec{c}\|$. তাহলে পরীক্ষা করে দেখি $\|\vec{a}\|^2 + \|\vec{b}\|^2 - \|\vec{c}\|^2$ কত হয়। যদি শূন্য হয়, তবে $\vec{a}, \vec{b}$ পরস্পরের সঙ্গে orthogonal হবে, নইলে নয়। হিসেব করলে দেখবে এরকম দাঁড়াচ্ছে--

$$\|\vec{a}\|^2 + \|\vec{b}\|^2 - \|\vec{c}\|^2 = \cdots = 2(a_1b_1 + a_2b_2).$$

তার মানে orthogonal হওয়া বা না হওয়া নির্ভর করছে $a_1b_1 + a_2b_2$-র শূন্য হওয়া বা না হওয়ার উপরে। এখানে $a_1b_1 + a_2b_2$ হল $\vec{a}, \vec{b}$-এর একটা function–

$$f(\vec{a}, \vec{b}) = a_1b_1 + a_2b_2.$$

একটু চিন্তা করলেই বুঝবে যে, একই কথা খাটে $\mathbb{R}^3$-এর ক্ষেত্রেও। সেখানে $f(\vec{a}, \vec{b})$ নিতে হবে এইভাবে--

$$f(\vec{a}, \vec{b}) = a_1b_1 + a_2b_2 + a_3b_3.$$

এবার যদি বলি $\mathbb{R}^4$-এর জন্য $f(\vec{a}, \vec{b})$ কীভাবে নেওয়া উচিত, অমনি তুমি ধাঁ করে বলে দেবে

$$f(\vec{a}, \vec{b}) = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4.$$

একইভাবে $\mathbb{R}^n$-এর জন্য নেবে

$$f(\vec{a}, \vec{b}) = a_1b_1 + \cdots + a_nb_n.$$

এই function-টা আমরা হায়ার সেকণ্ডারীতেই দেখেছি। তখন একে বলতাম dot product. Linear algebra-য় আমরা dot product-এর ধারণাটাকেই একটু generalise করে বলি **inner product**. অর্থাৎ dot product হল একটা বিশেষ ধরণের inner product. এই inner product থেকে আসবে norm-এর এই সংজ্ঞাটা--

$$\|\vec{a}\| = \sqrt{a_1^2 + \cdots + a_n^2}.$$

যদি কোনো দুটো vector-এর inner product শূন্য হয়, তবে তাদেরকে বলব পরস্পরের সঙ্গে orthogonal.
ঠিক কথা, কিন্তু $\mathbb{R}^4, \mathbb{R}^n$ ইত্যাদির বেলায় orthogonal হওয়া বা না হওয়া নিয়ে আদৌ মাথা ঘামাব কেন? ওসব ক্ষেত্রে তো আর ছবি আঁকা যাবে না! মনে রেখো যে $\mathbb{R}^2$ বা $\mathbb{R}^3$-র ক্ষেত্রে যে orthogonal হওয়া নিয়ে মাথা ঘামাচ্ছিলাম, সেটা কিন্তু ঠিক ছবি এঁকে হাতের সুখ করার জন্য নয়। একটা basis-এর vector-গুলো পরস্পরের সঙ্গে orthogonal হলে সেরকম basis নিয়ে কাজ করা সুবিধা হয়, সেইটাই ছিল আসল কারণ। এইরকম basis-দের বলে orthogonal basis. এদের নিয়ে কাজ করার সুবিধাটা কী, সেটা এবার একটু বুঝে নেওয়া যাক।

47.1 Orthogonal basis-এর প্রথম সুবিধা

Example 1: ধরো তোমাকে $\mathbb{R}^3$-র একটা basis দিলাম--

$$\{(1,2,3), \quad (2,3,-15), \quad (2,5,7)\}.$$

এবার তোমাকে একটা vector দিচ্ছি $(3,4,5)$. তুমি জানো যে, যেকোনো vector-কেই basis-এর linear combination হিসেবে প্রকাশ করা যায়, মানে এমন $a, b, c \in \mathbb{R}$ আছে যাতে

$$(3,4,5) = a(1,2,3) + b(2,3,-15) + c(2,5,7) \qquad (*)$$

হয়। তোমার কাজ হল a, b, c বার করা।

SOLUTION: এখানে a, b, c-র অস্তিত্ব নিয়ে সন্দেহমাত্র নেই, কিন্তু তা বলে সত্যি সত্যি হাতেকলমে ওদের বার করা বেশ বিরক্তিকর কাজ। তার জন্য আমাদের তিনটে equation বানাতে হবে a, b, c-কে unknown হিসেবে নিয়ে। যেমন (*)-এর প্রথম entry-গুলো নিয়ে হয় $a + 2b + 2c = 3$. এইভাবে পাওয়া যাচ্ছে--

$$\begin{aligned} a + 2b + 2c &= 3 \\ 2a + 3b + 5c &= 4 \\ 3a - 15b + 7c &= 5. \end{aligned}$$

এদের solve করা ভীষণ মজাদার কিছু নয়। যাই হোক আমরা solve করে পেয়েছি

$$a = \tfrac{33}{5}, \; b = \tfrac{1}{10}, \; c = -\tfrac{19}{10}.$$

■

Solve করে কী পাওয়া গেল, সেটা আমাদের কাছে বড় কথা নয়। বড় কথা হল-- এই solve করার কাজটা বেশ বিচ্ছিরি। এখানে আমরা $\mathbb{R}^3$ নিয়ে কাজ করছিলাম, যদি dimension আরও বেশী কিছু নিতাম, তবে equation এবং unknown-এর সংখ্যাও সেই হিসেবে বাড়ত, এবং কাজটা উত্তরোত্তর দুঃসাধ্য হয়ে উঠত। কিন্তু যদি basis-এর vector-গুলো পরস্পরের সঙ্গে orthogonal হত, তবে কাজটা জলের মত সোজা হয়ে যায়। কী করে হয়, সেটা বোঝার জন্য এই যে f বলে function-টার কথা বললাম--

$$f(\vec{x}, \vec{y}) = x_1y_1 + x_2y_2 + x_3y_3, \quad (\vec{x}, \vec{y} \in \mathbb{R}^3),$$

তার কিছু বৈশিষ্ট্য লক্ষ্য কর--

1. সর্বদাই $f(\vec{a}, \vec{b}) = f(\vec{b}, \vec{a})$ হয়।

2. $f(\vec{a}, \vec{b})$-র মধ্যে দুটো vector আছে $\vec{a}$ আর $\vec{b}$, যদি এদের মধ্যে যেকোনোটাকে কোনো সংখ্যা $\alpha \in \mathbb{R}$ দিয়ে গুণ কর তবে সেই α-টা f-এর বাইরে বেরিয়ে আসে--

$$f(\alpha\vec{a}, \vec{b}) = \alpha f(\vec{a}, \vec{b}).$$

যেহেতু $f(\vec{a}, \vec{b}) = f(\vec{b}, \vec{a})$, তাই α দিয়ে $\vec{b}$-কে গুণ করলেও একই ব্যাপার হত। এই "constant বেরিয়ে আসার" ব্যাপারটা অংকের অনেক জায়গাতেই দেখা যায়। তোমরা হায়ার সেকণ্ডারীতে differentiation-এর সময়েও দেখেছ যে

$$\frac{d}{dx}(\alpha f(x)) = \alpha \frac{d}{dx} f(x)$$

হয়। Integration-এর বেলাতেও একইরকম।

3. আমাদের $f(\vec{a}, \vec{b})$-এর আরেকটা গুণ আছে। যদি তুমি $\vec{a}$-এর জায়গায় $\vec{a}_1 + \vec{a}_2$ নাও, তবে

$$f(\vec{a}_1 + \vec{a}_2, \vec{b}) = f(\vec{a}_1, \vec{b}) + f(\vec{a}_2, \vec{b})$$

হবে। যেহেতু $f(\vec{a}, \vec{b}) = f(\vec{b}, \vec{a})$, তাই $\vec{b}$-কেও একইভাবে ভাঙা যায়--

$$f(\vec{a}, \vec{b}_1 + \vec{b}_2) = f(\vec{a}, \vec{b}_1) + f(\vec{a}, \vec{b}_2).$$

এটাও হায়ার সেকণ্ডারীতে differentiation বা integration-এর বেলায় আমরা দেখেছি--

$$\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx} f(x) + \frac{d}{dx} g(x).$$

এইবার আমরা দেখব যে, এই বৈশিষ্ট্যগুলোর দৌলতে আমাদের জীবন কত সহজ হয়ে যাবে। ধরো তোমাকে একটা basis দিলাম--

$$\{\vec{v}_1, ..., \vec{v}_n\},$$

যেখানে $\vec{v}_i$-রা সবাই নিজেদের মধ্যে orthogonal. একটা উদাহরণ মাথায় রাখলে সুবিধা হতে পারে-- $\mathbb{R}^3$-এর মধ্যে

$$\vec{v}_1 = (1, 2, 3), \quad \vec{v}_2 = (2, -1, 0), \quad \vec{v}_3 = (3, 6, -5).$$

এরা যে সত্যিই orthogonal, সেটা একটু পরীক্ষা করে দেখলেই বুঝবে, যেমন

$$f(\vec{v}_1, \vec{v}_2) = 1 \times 2 + 2 \times (-1) + 3 \times 0 = 0.$$

এবার তোমাকে একটা vector দিচ্ছি $\vec{v} = (3, 4, 5)$. তোমার কাজ হল $\vec{v}$-কে এই basis-টার linear combination হিসেবে লেখা। এখানেও অবশ্যই তুমি আবার equation solve করার জটিল পথে এগোতে পারো। কিন্তু অনেক সহজ কায়দা হল এইভাবে এগোনো--আমরা জানি যে এমন a, b, c আছে যাতে

$$\vec{v} = a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3$$

হয়। এবার আমরা $f(\vec{v}, \vec{v}_1)$ বার করব--

$$\begin{aligned} f(\vec{v}_1, \vec{v}) &= f(\vec{v}_1, a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3) \\ &= f(\vec{v}_1, a\vec{v}_1) + f(\vec{v}_1, b\vec{v}_2) + f(\vec{v}_1, c\vec{v}_3) \quad [\because \text{ভেঙে যায়}] \\ &= af(\vec{v}_1, \vec{v}_1) + bf(\vec{v}_1, \vec{v}_2) + cf(\vec{v}_1, \vec{v}_3) \quad [\because \text{constant বেরিয়ে আসে}] \\ &= af(\vec{v}_1, \vec{v}_1). \end{aligned}$$

সুতরাং আমরা পেয়ে গেলাম

$$a = \frac{f(\vec{v}_1, \vec{v})}{f(\vec{v}_1, \vec{v}_1)}.$$

এখানে অবশ্য তুমি আপত্তি তুলতে পারো এই বলে যে $f(\vec{v}_1, \vec{v}_1)$ যে 0 হয়ে যাবে না, তার কোনো গ্যারান্টী আছে কি না। অবশ্যই আছে, কারণ

$$f(\vec{v}_1, \vec{v}_1) = 1^2 + 2^2 + 3^2,$$

যেটা আসলে $\vec{v}_1$-এর দৈর্ঘ্যের square ছাড়া কিছুই নয়, সুতরাং সেটা ≥ 0 হতে বাধ্য। এটা 0 হতে পারে খালি $\vec{v}_1 = \vec{0}$ হলে, যেটা এখানে অসম্ভব, কারণ হাজার হোক $\vec{v}_1$ হল একটা basis-এর সদস্য!
একইভাবে $f(\vec{v}_2, \vec{v})$ আর $f(\vec{v}_3, \vec{v})$ নিয়ে কাজ করলে যে

$$b = \frac{f(\vec{v}_2, \vec{v})}{f(\vec{v}_2, \vec{v}_2)}, \quad c = \frac{f(\vec{v}_3, \vec{v})}{f(\vec{v}_3, \vec{v}_3)}$$

পাবে, সেটা অংকটা না কষেই আন্দাজ করা যায়। সুতরাং এখানে কোনো জটিল equation solve করার দরকারই হল না।

Exercise 1: $\mathbb{R}^4$-এর এই একটা orthogonal basis দিলাম $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$, যেখানে--

$$\begin{aligned} \vec{v}_1 = (1, 1, 1, 1) \quad &, \quad \vec{v}_2 = (1, -1, 1, -1), \\ \vec{v}_3 = (1, 1, -1, -1) \quad &, \quad \vec{v}_4 = (1, -1, -1, 1). \end{aligned}$$

এবার একটা vector দিচ্ছি $\vec{v} = (1, 2, -3, 4)$. কোনো equation solve না করে $\vec{v}$-কে এই orthogonal basis-এর একটা linear combination হিসেবে প্রকাশ কর দেখি। ■

এই হল প্রথম সুবিধা। এবার দ্বিতীয় আরেকটা সুবিধার কথা বলি।

47.2 Orthogonal basis-এর দ্বিতীয় সুবিধা

একটু আগে তোমাকে একটা অংক দিয়েছিলাম, যেখানে কাজ ছিল $(3,4,5)$-কে এই basis-টার linear combination হিসেবে লেখা--

$$\{(1,2,3), \quad (2,3,-15), \quad (2,5,7)\}.$$

আমরা দেখেছিলাম যে, কাজটা একটা বিচ্ছিরি equation solve করার কাজে পরিণত হয়। পুরোটা solve করে পেয়েছিলাম--

$$a = \tfrac{33}{5}, \;\; b = \frac{1}{10}, \;\; c = -\tfrac{19}{10}.$$

এবার নীচের অংকটা দ্যাখো।

Example 2: আবার তোমাকে একটা basis দিলাম

$$\{(1,2,3), \quad (2,3,-15), \quad (2,5,17)\}.$$

এটা প্রায় আগের basis-টাই, খালি তৃতীয় vector-টাকে সামান্য পালটে দিয়েছি। আবারও সেই vector-টাই নিলাম--

$$(3,4,5).$$

তোমার কাজ হল একে এই basis-এর linear combination হিসেবে লেখা।
SOLUTION: আবার যদি আগের মতই equation solve করি, তবে পাব--

$$(3,4,5) = a(1,2,3) + b(2,3,-15) + c(2,5,17),$$

যেখানে

$$a = \tfrac{71}{5}, \quad b = -\tfrac{9}{5}, \quad c = -\tfrac{19}{5}.$$

এখানে বিস্ময়কর কথা হল a এবং b-টা আগের থেকে বদলে গেছে, যদিও আমাদের basis-এর প্রথম দুটো vector একই আছে! ■

এই ব্যাপারটা একটু তলিয়ে ভাবা যাক। আগের অংকটায় যখন লিখেছিলাম

$$(3,4,5) = a(1,2,3) + b(2,3,-15) + c(2,5,7),$$

তখন এমন ভেবে ফেলা স্বাভাবিক যে, a, b, c হল $(3,4,5)$-এর মধ্যে যথাক্রমে $(1,2,3), (2,3,-15)$ এবং $(2,5,7)$-এর অবদান। সেখানে আমরা $a = \frac{33}{5}$ পেয়েছিলাম, সুতরাং সেইভাবে ভাবলে $(3,4,5)$-এর মধ্যে $(1,2,3)$-এর অবদানের পরিমাণ হল $\frac{33}{5}$. কিন্তু তাই যদি ভাবি, তবে এটাও ভাবতে হয় যে, পরের অংকে $(3,4,5)$-এর মধ্যে সেই একই $(1,2,3)$-এর অবদানের পরিমাণ হল $\frac{71}{5}$. দুটো vector সেই একই রইল, অথচ অবদানের পরিমাণটা গেল বদলে, এ আবার কেমন কথা? আসলে এই অবদানের গল্পটা ভাবতে স্বাভাবিক মনে হলেও ওর সত্যি কোনো মানেই হয় না।
কিন্তু যদি আমাদের basis-এর vector-গুলো পরস্পরের সঙ্গে orthogonal হত, তবে কিন্তু আর সে সমস্যা থাকত না। কারণটা বোঝা খুবই সহজ--যদি $\{\vec{v}_1, ..., \vec{v}_n\}$ একটা orthogonal basis হয়, তবে যেকোনো $\vec{v}$-কে আমরা

$$\vec{v} = \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n$$

হিসেবে লিখে ফেলতে পারি, যেখানে--

$$\alpha_i = \frac{f(\vec{v}_i, \vec{v})}{f(\vec{v}_i, \vec{v}_i)}.$$

সুতরাং α_i সম্পূর্ণভাবেই $\vec{v}$ এবং $\vec{v}_i$ দ্বারা নির্ধারিত, basis-এর বাকি সদস্যদের কোনো প্রভাব সেখানে নেই। সুতরাং এখানে আমরা সেই স্বাভাবিক কথাটা সত্যিই নিশ্চিন্তে বলতে পারি--α_i হল $\vec{v}$-এর মধ্যে $\vec{v}_i$-এর অবদান।

47.3 Matrix-এর গুণ দিয়ে dot product লেখা

এই যে function-টাকে নিয়ে আমরা কাজ করছি, $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, যেখানে

$$f(\vec{x}, \vec{y}) = x_1 y_1 + \cdots + x_n y_n,$$

এটার নাম যে dot product, সে তো জানোই। এটাকে matrix multiplication দিয়ে লেখা যায়। তার জন্য $\vec{x}$ আর $\vec{y}$ দুজনকেই column vector হিসেবে কল্পনা করো, মানে $n \times 1$ সাইজের matrix হিসেবে। তবে--

$$f(\vec{x}, \vec{y}) = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \vec{x}'\vec{y}.$$

এখানে $\vec{x}$-এর transpose নিয়ে ওকে শুইয়ে দিয়েছি। চাইলে তুমি $\vec{y}$-কেও শোয়াতে পারতে। তবে পেতে--

$$f(\vec{x}, \vec{y}) = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \vec{y}'\vec{x}.$$

যাকেই শোয়াও না কেন, শোয়ানো vector-টা থাকবে বাঁদিকে, খাড়া vector-টা থাকবে ডানদিকে।
এখানে আমরা দুটো vector-এর মধ্যে গুণটা করছি। একই ব্যাপার হয় যেকোনো দুটো matrix-এর মধ্যে গুণের বেলাতেও। ধরো $A_{m\times n}$ আর $B_{n\times p}$ হল দুটো matrix. তবে AB হবে একটা $m \times p$ সাইজের matrix. এর (i, j)-th entry হবে

$$\sum_{k=1}^{n} a_{ik} b_{kj},$$

যেটা A-র i-th row এবং B-র j-th column-এর dot product ছাড়া আর কিছুই নয়। এই ব্যাপারটা ভালো করে চিন্তা করে বুঝে নাও। বিশেষ করে যদি কোনো একটা $\vec{x}$-এর সাথে তুমি অনেকগুলো vector, ধরো $\vec{a}_1, ..., \vec{a}_k$-র dot product বার করতে চাও, তবে $\vec{a}_i$-গুলো একটার উপর একটা শুইয়ে একটা $A_{k\times n}$ বানিয়ে রাখো, এবং $A\vec{x}$ বার করো। এর entry-গুলোই হবে $f(\vec{a}_i, \vec{x})$. এই ব্যাপারটা এই অধ্যায়ে আমাদের পরে কাজ আসবে।

Exercise 2: এমন একটা matrix আছে, $A_{n\times n}$, যাতে $A'A$ একটা diagonal matrix হয়। যদি A-র i-th column-টাকে $\vec{a}_i$ নাম দিই, তবে $i \neq j$ হলে $\vec{a}_i$ এবং $\vec{a}_j$-র dot product কত হবে? ■

Exercise 3: If $\vec{v} \in \mathbb{R}^n$ is such that the dot product between $\vec{v}$ and $2\vec{v}$ is zero, find $\vec{v}$. ■

DAY 48
Inner product (part 2)

48.1 Dot product থেকে inner product

আমরা এতক্ষণ dot product নিয়ে কাজ করছিলাম--

$$f(\vec{x}, \vec{y}) = x_1 y_1 + \cdots + x_n y_n.$$

এর উপযোগিতার পিছনে কাজ করছিল এই কয়টা বৈশিষ্ট্য--

1. $\forall \vec{x} \in V \quad f(\vec{x}, \vec{x}) \geq 0$ এবং $f(\vec{x}, \vec{x}) = 0$ হবে একমাত্র তখনই যখন $\vec{x} = \vec{0}$.
2. $\forall \vec{x}, \vec{y} \in V \quad f(\vec{x}, \vec{y}) = f(\vec{y}, \vec{x})$.
3. "ভেঙে যাওয়া" এবং "constant বেরিয়ে আসা", মানে--

$$\forall \vec{x}_1, \vec{x}_2, \vec{y} \in V \quad \forall \alpha, \beta \in \mathbb{R} \quad f(\alpha\vec{x}_1 + \beta\vec{x}_2, \vec{y}) = \alpha f(\vec{x}_1, \vec{y}) + \beta f(\vec{x}_2, \vec{y}).$$

যেহেতু $f(\vec{x}, \vec{y}) = f(\vec{y}, \vec{x})$, তাই

$$\forall \vec{x}, \vec{y}_1, \vec{y}_2, \in V \quad \forall \alpha, \beta \in \mathbb{R} \quad f(\vec{x}, \alpha\vec{y}_1 + \beta\vec{y}_2) = \alpha f(\vec{x}, \vec{y}_1) + \beta f(\vec{x}, \vec{y}_2).$$

Dot product এমন একটা function, যেটা এই শর্তগুলো পালন করে। কিন্তু dot product-ই এরকম একমাত্র function নয়। যেসব $f(\vec{a}, \vec{b})$ এই শর্তগুলো পালন করে, তাদেরকে অংকের ভাষায় বলে **inner product**. লেখার সময়ে সাধারণতঃ $f(\vec{a}, \vec{b})$ না লিখে $\left\langle \vec{a}, \vec{b} \right\rangle$ লেখে। যেমন যদি $\mathbb{R}^2$-এর বেলায়

$$\left\langle \vec{a}, \vec{b} \right\rangle = a_1 b_1 + 2a_2 b_2$$

নিতাম, তাহলেও এই সবগুলো ধর্মই পালিত হত (পরীক্ষা করেই দেখো!)। তাই এটাও একটা inner product.
বস্তুতঃ এই বৈশিষ্ট্যগুলো আমাদের ছোটোবেলাকার পরিচিত গুণ করার বিভিন্ন ধর্মের মতই। খালি $\left\langle \vec{a}, \vec{b} \right\rangle$-এর বদলে ab কল্পনা কর যেখানে $a, b \in \mathbb{R}$. তাহলে প্রথম শর্তটা বলছে যেকোনো a-র জন্যই $a^2 \geq 0$ এবং $a^2 = 0$ একমাত্র তখনই হবে, যখন $a = 0$. সেটা তো আমরা ছোটোবেলা থেকেই জানি! দ্বিতীয় শর্তটা আসলে গুণের commutativity, মানে $ab = ba$. আর তৃতীয় শর্তটা হল যোগের উপরে গুণের distributivity.
এখানে বলে রাখি যে, সব vector space-এর ক্ষেত্রেই এরকম f পাওয়া যাবে না। খালি যখন field-টা $\mathbb{R}$ বা $\mathbb{C}$ হবে, খালি তখনই inner product নিয়ে আলোচনা করা হয়। একটা vector space-এর উপর যখন একটা inner product থাকে, তখন inner product-সুদ্ধ vector space-টাকে বলে একটা **inner product space**. যদি কখনো লেখা দেখো real inner product space, তার মানে field-টা হল $\mathbb{R}$. একইভাবে field-টা $\mathbb{C}$ হলে বলতাম complex inner product space.

DEFINITION: **(Real) Inner product**

By an **inner product** on a real vector space V we mean a function

$$f : V \times V \to \mathbb{R}$$

such that

1. $\forall \vec{v} \in V \quad f(\vec{v}, \vec{v}) \geq 0$, and $= 0 \iff \vec{v} = \vec{0}$,
2. $\forall \vec{u}, \vec{v} \in V \quad f(\vec{u}, \vec{v}) = f(\vec{v}, \vec{u})$,
3. $\forall \vec{u}, \vec{v}, \vec{w} \in V \quad \forall a, b \in \mathbb{R}$

$$f(a\vec{u} + b\vec{v}, \vec{w}) = af(\vec{u}, \vec{w}) + bf(\vec{v}, \vec{w}).$$

Usually we write $f(\vec{u}, \vec{v})$ as $\langle \vec{u}, \vec{v} \rangle$.

যদি field-টা $\mathbb{C}$ হত, তবে সংজ্ঞাটা সামান্য পাল্টাতে হত--

> **DEFINITION: (Complex) Inner product**
>
> By an **inner product** on a complex vector space V we mean a function $f : V \times V \to \mathbb{C}$ such that
>
> 1. $\forall \vec{v} \in V \quad f(\vec{v}, \vec{v})$ is a real number ≥ 0, and $= 0 \iff \vec{v} = \vec{0}$,
> 2. $\forall \vec{u}, \vec{v} \in V \quad f(\vec{u}, \vec{v}) = \overline{f(\vec{v}, \vec{u})}$,
> 3. $\forall \vec{u}, \vec{v}, \vec{w} \in V \quad \forall a, b \in \mathbb{C}$
>
> $$f(a\vec{u} + b\vec{v}, \vec{w}) = af(\vec{u}, \vec{w}) + bf(\vec{v}, \vec{w}).$$

$\mathbb{C}$-এর বেলায় কেন এই সামান্য পরিবর্তনটুকুর দরকার পড়ল, সেটা বোঝার জন্য লক্ষ্য কর কীভাবে আমরা inner product ব্যবহার করে কোনো vector-এর norm (দৈর্ঘ্য) বার করি। ধরো $\vec{v} = (x, y, z) \in \mathbb{R}^3$. তবে

$$\|\vec{v}\| = \sqrt{x^2 + y^2 + z^2} = \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

এটাকেও সাধারণ গুণের উপমা দিয়ে ভাবা যায়--কোনো $x \in \mathbb{R}$ দেওয়া থাকলে 0 থেকে তার দূরত্ব হল $|x| = \sqrt{x^2}$. যদি $z \in \mathbb{C}$ নিয়ে কাজ করি, তবে কিন্তু এটা খাটে না। তখন যেটা খাটে সেটা হল $|z| = \sqrt{z\bar{z}}$. লক্ষ্য কর এখানে কিভাবে z-এর complex conjugate এসে পড়ল। ঠিক একই কারণে complex inner product-এর সংজ্ঞায় complex conjugate ব্যবহার করতে হয়।
লক্ষণীয় যে, real inner product-কে দুদিক থেকেই ভেঙে constant বার করা যায়, মানে

$$\langle a\vec{x}_1 + b\vec{x}_2, \vec{y}_1 \rangle = a \langle \vec{x}_1, \vec{y} \rangle + b \langle \vec{x}_2, \vec{y} \rangle \text{ এবং } \langle \vec{x}, a\vec{y}_1 + b\vec{y}_2 \rangle = a \langle \vec{x}, \vec{y}_1 \rangle + b \langle \vec{x}, \vec{y}_2 \rangle$$

হয়। কিন্তু complex inner product-এর বেলায় কাজটা খালি বাঁদিক থেকে করা যায়। ডানদিকে করতে গেলে constant-গুলো complex conjugate হয়ে যায়, কারণ--

$$\begin{aligned} \langle \vec{z}, a\vec{w}_1 + b\vec{w}_2 \rangle &= \overline{\langle a\vec{w}_1 + b\vec{w}_2, \vec{z} \rangle} \\ &= \bar{a}\overline{\langle \vec{w}_1, \vec{z} \rangle} + \bar{b}\overline{\langle \vec{w}_2, \vec{z} \rangle} \\ &= \bar{a} \langle \vec{z}, \vec{w}_1 \rangle + \bar{b} \langle \vec{z}, \vec{w}_2 \rangle . \end{aligned}$$

যেকোনো inner product-এর জন্যই তা থেকে একটা norm পাওয়া যায়। তাকে অনেক সময়ে বলে **norm induced by inner product**.

> **DEFINITION: Norm induced by inner product**
>
> Let V be any inner product space with inner product given by $\langle \cdot, \cdot \rangle$. Then for any $\vec{v} \in V$ its **norm** is defined as
> $$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

যেহেতু $\langle \vec{v}, \vec{v} \rangle \geq 0$, তাই square root নিতে কোনো অসুবিধা হবে না।
এবার সংজ্ঞার শর্তগুলো নিয়ে একটু সড়গড় হয়ে নেওয়া যাক। সেই জন্যই নীচের কয়েকটা অংক।

Exercise 4: দেখাও যে, $\mathbb{R}^4$-এর উপরে এইটা একটা inner product হয়--

$$\langle \vec{x}, \vec{y} \rangle = 10x_1y_1 + 3x_2y_2 + \frac{1}{2}x_3y_3 + x_4y_4,$$

যেখানে $\vec{x} = (x_1, ..., x_4)$ এবং $\vec{y} = (y_1, ..., y_4)$.
HINT: (Real) inner product-এর শর্ত তিনটে সরাসরি পরীক্ষা করে গেলেই হবে। ■

Exercise 5: দেখাও যে, $\mathbb{R}^4$-এর উপরে এইটা একটা inner product হয় না--

$$\langle \vec{x}, \vec{y} \rangle = 10x_1y_1 - 3x_2y_2 + \frac{1}{2}x_3y_3 + x_4y_4.$$

HINT: এখানে সমস্যা হবে ওই -3-টা নিয়ে। ■

এই দুটো অংকে থেকে যা শিখলে সেটা মেলালেই পাবে নীচের অংকটা।

Exercise 6: Show that

$$\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} a_i x_i y_i$$

is an inner product on $\mathbb{R}^n$ if and only if $\forall i \quad a_i > 0$.
■

Exercise 7: Is $f(\vec{x}, \vec{y}) = x_1y_1$ an inner product on $\mathbb{R}^2$? ■

Example 3: In $\mathbb{R}^3$ let $\alpha = (a_1, a_2, a_3), \beta = (b_1, b_2, b_3)$ and

$$\langle \alpha, \beta \rangle = |a_1b_1 + a_2b_2 + a_3b_3|.$$

Does $\langle \alpha, \beta \rangle$ define a real inner product? Answer with reason.
SOLUTION: এই রকম অংকে প্রথমে মনে মনে ভেবে নিতে হয় inner product-এর তিনটে শর্তই পালিত হচ্ছে কিনা। যদি হয়, তবে তিনটেকেই প্রমাণ করে দেখাতে হবে। যদি একটাও পালিত না হয়, তবে খালি সেটুকু উল্লেখ করে তার জন্য একটা counterexample দিলেই হবে, যেমন এইক্ষেত্রে ওই absolute value-র ব্যাপারটা দেখেই সন্দেহ হচ্ছে যে, constant বেরিয়ে আসতে সমস্যা হবে। যদি constant-টা negative হয়, তবে ওই absolute value-র ধাক্কায় বেচারী তার চিহ্নটা খোয়াবে। তাই counterexample বানানোর জন্য কোনো negative constant দিয়ে গুণ করব, ধরো -2 দিয়ে--

> No, it is not a real inner product. For example, if $\alpha = \beta = (1, 1, 1)$ we have
>
> $$\begin{aligned} \langle -2\alpha, \beta \rangle &= \langle (-2, -2, -2), (1, 1, 1) \rangle \\ &= |(-2) \times 1 + (-2) \times 1 + (-2) \times 1| \\ &= 6. \end{aligned}$$

But

$$\begin{aligned}\langle \alpha, \beta\rangle &= \langle (1,1,1),(1,1,1)\rangle \\ &= |1\times 1+1\times 1+1\times 1| \\ &= 3.\end{aligned}$$

তার মানে এখানে $\langle -2\alpha, \beta\rangle$ মোটেই $-2\langle \alpha, \beta\rangle$-এর সমান হচ্ছে না--

So we have

$$\langle -2\alpha, \beta\rangle = 6 \neq -2\times 3 = -2\langle \alpha, \beta\rangle.$$

■

Exercise 8: In $\mathbb{R}^3$ let $\vec{\alpha} = (a_1, a_2, a_3)$, $\vec{\beta} = (b_1, b_2, b_3)$. Determine whether $\langle , \rangle$ is a real inner product for $\mathbb{R}^3$ if $\langle , \rangle$ is defined by

$$\left\langle \vec{\alpha}, \vec{\beta}\right\rangle = a_1 b_1 + (a_2 + a_3)(b_2 + b_3).$$

HINT:
$\vec{\alpha} = (0, 1, -1) \neq \vec{0}$ নিয়ে $\langle \vec{\alpha}, \vec{\alpha}\rangle$ বার করে দেখো তো > 0 হয় কিনা! ■

Inner product নানা রকম হয়। তার মধ্যে কয়েকটা দেখতে বেশ অদ্ভুত। তবে সবচেয়ে বেশী কাজে লাগে dot product,

$$x_1 y_1 + \cdots + x_n y_n,$$

যার পোশাকী নাম হল Euclidean (ইউক্লিডিয়ান) inner product. এই নামের কারণ হল $n = 2$ বা 3 নিলে এই inner product অনুযায়ী orthogonal আর জ্যামিতির perpendicular সমার্থক। আর গ্রীক জ্যামিতির জনক হিসেবে Euclid-এর নাম কে না জানে!
অনেক সময়ে Euclidean space বলে একটা কথা শুনবে। এর মানে হল $\mathbb{R}$-কে field নিয়ে যেকোনো finite dimensional inner product space. এর মধ্যে গুলিয়ে যাবার বিষয়টা হল--Euclidean space-এর inner product-টা যে Euclidean inner product হবেই, এমন কোনো কথা নেই!

Example 4: Find the norm of the vector $u = (2, 1, 1)$ in Euclidean space $\mathbb{R}^3$ with respect to the usual inner product.
SOLUTION:

$$\|u\| = \sqrt{\langle u, u\rangle} = \sqrt{2^2 + 1^2 + 1^2} = \sqrt{6}.$$

■

48.2 Euclidean inner product-এর complex সংস্করণ

যদি আমরা $\mathbb{R}^n$ নিয়ে কাজ করি, তবে

$$\langle \vec{x}, \vec{y}\rangle = \sum_{i=1}^{n} x_i y_i$$

একটা inner product হবে, কিন্তু $\mathbb{C}^n$-এর বেলায় তা হবে না। যেমন $\mathbb{C}^2$-এর বেলায় এই সংজ্ঞা লাগালে (i,i)-এর norm হয়ে যাবে $i^2+i^2=-2$-এর square root, যেটা কোনো মতেই একটা nonnegative real সংখ্যা নয়।
তবে একটু মেরামতি করে নিলেই $\mathbb{C}^n$-এর জন্যও একটা inner product পাওয়া যাবে--

$$\langle \vec{w}, \vec{z} \rangle = \sum_{i=1}^{n} w_i \bar{z}_i.$$

এখানে একটা ব্যাপারে সতর্ক থেকো। আমরা z_i-গুলোর complex conjugate নিয়েছি, w_i-গুলোর নয়। যদি w_i-গুলো complex conjugate নিতে, তবে কিন্তু inner product হত না (কেন?)। এই ব্যাপারটা আরো প্রকট হয়ে ওঠে matrix-এর গুণ দিয়ে লিখতে গেলে। $\mathbb{R}^n$-এর বেলায় আমরা দেখেছি Euclidean inner product-কে দুইভাবে লেখা যায়--

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}'\vec{y} = \vec{y}'\vec{x}.$$

কিন্তু $\mathbb{C}^n$-এর বেলায় জোর দিয়ে খালি বলা যায়--

$$\langle \vec{w}, \vec{z} \rangle = \vec{z}^*\vec{w},$$

যেটা সবসময়ে $\vec{w}^*\vec{z}$-এর সমান নাও হতে পারে। এই জায়গাটায় সাবধান। যখন $\langle \vec{w}, \vec{z} \rangle$ লিখছি, তখন $\vec{w}$ রয়েছে বাঁদিকে। কিন্তু একই জিনিস যখন $\vec{z}^*\vec{w}$ আকারে লিখছি, তখন $\vec{w}$ চলে যাচ্ছে ডানদিকে!

48.3 কিছু সহজ অংক

Example 5: Prove that in a real inner product space V

$$\|c\alpha\| = |c| \;\; \|\alpha\|$$

for $\alpha \in V$ and $c \in \mathbb{R}$.
SOLUTION: ছোট্টো অংক, বেশী কিছু করার নেই, ধাপগুলো গুছিয়ে লিখি--

$$\begin{aligned}
\|c\alpha\| &= \sqrt{\langle c\alpha, c\alpha \rangle} \\
&= \sqrt{c\,\langle \alpha, c\alpha \rangle} \quad [\because \langle a\vec{u}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle] \\
&= \sqrt{c\,\langle c\alpha, \alpha \rangle} \quad [\because \langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle] \\
&= \sqrt{c^2\,\langle \alpha, \alpha \rangle} \quad [\because \langle a\vec{u}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle] \\
&= \sqrt{c^2}\sqrt{\langle \alpha, \alpha \rangle} \\
&= |c| \;\; \|\alpha\|,
\end{aligned}$$

as required.

■

Example 6: Prove that, for all α, β in a Euclidean space V,

$$\langle \alpha+\beta, \alpha-\beta \rangle = 0 \text{ if and only if } \|\alpha\| = \|\beta\|.$$

[2]
SOLUTION: যদি inner product-কে গুণের উপমা দিয়ে ভাবো, তবে এই অংকের ধাপগুলো অনায়াসেই দেখতে পাবে। যদি $(\alpha+\beta)(\alpha-\beta)=0$ হয়, তবে $\alpha^2-\beta^2=0$. সুতরাং $|\alpha|=|\beta|$. এই জিনিসটা মাথায় রেখে এবার সমাধানের ধাপগুলো দেখো।

We have

$$\begin{aligned}\langle \alpha+\beta, \alpha-\beta\rangle &= \langle \alpha, \alpha-\beta\rangle + \langle \beta, \alpha-\beta\rangle \\ &= \langle \alpha, \alpha\rangle - \langle \alpha, \beta\rangle + \langle \beta, \alpha\rangle - \langle \beta, \beta\rangle \\ &= \langle \alpha, \alpha\rangle - \langle \beta, \beta\rangle \quad [\because \langle \alpha, \beta\rangle = \langle \beta, \alpha\rangle] \\ &= \|\alpha\|^2 - \|\beta\|^2 .\end{aligned}$$

Thus

$$\begin{aligned}& \langle \alpha+\beta, \alpha-\beta\rangle = 0 \\ \iff & \|\alpha\|^2 = \|\beta\|^2 \\ \iff & \|\alpha\| = \|\beta\| ,\end{aligned}$$

as required, since norm is always ≥ 0.

■

Example 7: For any two vectors α, β of a Euclidean space, show that $\langle \alpha, \beta\rangle = 0$ if and only if

$$\|\alpha+\beta\| = \|\alpha-\beta\| ,$$

where $\|\alpha\|$ denotes the norm of α.

SOLUTION:

We have

$$\begin{aligned}\|\alpha+\beta\| &= \sqrt{\langle \alpha+\beta, \alpha+\beta\rangle} \\ &= \sqrt{\langle \alpha, \alpha\rangle + \langle \alpha, \beta\rangle + \langle \beta, \alpha\rangle + \langle \beta, \beta\rangle} \\ &= \sqrt{\langle \alpha, \alpha\rangle + 2\langle \alpha, \beta\rangle + \langle \beta, \beta\rangle} \quad [\because \langle \alpha, \beta\rangle = \langle \beta, \alpha\rangle]\end{aligned}$$

Similarly,

$$\|\alpha-\beta\| = \sqrt{\langle \alpha, \alpha\rangle - 2\langle \alpha, \beta\rangle + \langle \beta, \beta\rangle}.$$

Thus

$$\begin{aligned}& \|\alpha+\beta\| = \|\alpha-\beta\| \\ \iff & (\|\alpha+\beta\|)^2 = (\|\alpha-\beta\|)^2 \\ \iff & \langle \alpha, \alpha\rangle + 2\langle \alpha, \beta\rangle + \langle \beta, \beta\rangle = \langle \alpha, \alpha\rangle - 2\langle \alpha, \beta\rangle + \langle \beta, \beta\rangle \\ \iff & \langle \alpha, \beta\rangle = -\langle \alpha, \beta\rangle \\ \iff & \langle \alpha, \beta\rangle = 0,\end{aligned}$$

as required.

■

Exercise 9: আচ্ছা, যদি উপরের অংকে field-টা $\mathbb{C}$ হত, তবেও কি চলত? ■

Exercise 10: If α, β are two orthogonal vectors in a Euclidean space V, then show that

$$\|\alpha + \beta\|^2 = \|\alpha - \beta\|^2.$$

HINT:
আগের উদাহরণটাই, এবার খালি একটা দিক করতে হবে। ■

নীচের অংকটাও একইরকম, তোমার হাত পাকানোর জন্য রেখে দিলাম।

Exercise 11: Prove that for vectors α, β in a Euclidean space, $\langle \alpha, \beta \rangle = 0$ if and only if $\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2$.
HINT: এই অংকটা আমরা অধ্যায়ের গোড়াতেই দেখেছিলাম Pythagoras's theorem-এর প্রসঙ্গে। ■

Example 8: Let β be a basis for a finite-dimensional inner product space. Prove that if $\langle x, z \rangle = 0$ for all $z \in \beta$, then $x = 0$.[3] **(2020.CC12.3di)**
SOLUTION:

Let the basis be $\{\vec{v}_1, ..., \vec{v}_n\}$.
Given: $\forall i \quad \langle \vec{x}, \vec{v}_i \rangle = 0$.
Let $\vec{x} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$.
Then $\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle = \langle \vec{x}, \sum_i a_i\vec{v}_i \rangle = \overline{\langle \sum_i a_i\vec{v}_i, \vec{x} \rangle} = \sum \bar{a}_i \overline{\langle \vec{v}_i, \vec{x} \rangle} = \sum \bar{a}_i \langle \vec{x}, \vec{v}_i \rangle = 0.$

লক্ষ্য করো, এখানে complex conjugate নিয়েছি, কারণ field-টা কী বলে দেয় নি। তাই $\mathbb{C}$ নিয়ে কাজ করেছি। যদি field-টা $\mathbb{R}$ হয়, তাহলেও সমাধানটা ভুল নয়, কারণ কোনো real সংখ্যার complex conjugate নিলে সংখ্যাটা অপরিবর্তিত থাকে।

Hence $\vec{x} = \vec{0}$, as required.

■

Exercise 12: Show that the sum of two inner products is again an inner product.[2] **(2020.CC12.3dii)**
HINT: সংজ্ঞাটা লাগালেই মুহূর্তের মধ্যে হয়ে যাবে। ■

এবারের অংকটা দেখতে একটু জটিল, এবং একটা কৌশল না জানলে করা শক্ত।

Example 9: Show that

$$\left\langle \sum_j a_j x^j, \sum_k b_k x^k \right\rangle = \sum_{j,k} \frac{a_j b_k}{j+k+1}$$

defines an inner product on the space $\mathbb{R}[x]$ of polynomials over the field $\mathbb{R}$.[5] **(2021.CC12.3f)**

SOLUTION: এখানে notation-টা বুঝে নাও। আমরা কাজ করছি polynomial-দের নিয়ে। $\sum_j a_j x^j$ দিয়ে একটা polynomial বোঝাচ্ছি। এই sum-এ j কোথা থেকে কোথায় যাচ্ছে, কিছু বলে দেয় নি। আমরা জানি যে, একটা polynomial-এ finite-সংখ্যক term থাকে, যাদের প্রত্যেকেরই চেহারা $a_j x^j$-র মত। যতগুলো এরকম term, ঠিক ততগুলো জিনিস sum-টাতে যোগ করা হচ্ছে। এরকম যাবতীয় polynomial-দের নিয়ে যে vector space-টা হয়, সেটাকেই লেখা হয়েছে $\mathbb{R}[x]$.

এই অংকে আমাদের কাজ হল inner product-এর সংজ্ঞার শর্ত তিনটে পরপর পরীক্ষা করে যাওয়া। দ্বিতীয় আর তৃতীয় শর্ত পরীক্ষা করা খুবই সহজ (তোমার জন্য রেখে দিলাম)। প্রথম শর্তটা দেখাতেই কৌশল লাগবে। সেটুকুই করে দেখাচ্ছি--

> We see that
>
> $$\left\langle \sum_j a_j x^j, \sum_k b_k x^k \right\rangle = \sum_{j,k} \frac{a_j b_k}{j+k+1} = \int_0^1 \left(\sum_j a_j x^j\right)\left(\sum_k b_k x^k\right) dx.$$

এটা কেন হচ্ছে, সেটা দেখা খুবই সহজ, কারণ $\int_0^1 x^{j+k}\,dx = \frac{1}{j+k+1}$.

> Thus, for all $p(x), q(x) \in \mathbb{R}[x]$ we have $\langle p(x), q(x)\rangle = \int_0^1 p(x)q(x)\,dx$.
> So, for all $p(x) \in \mathbb{R}[x]$ we have $\langle p(x), p(x)\rangle = \int_0^1 p(x)^2\,dx \geq 0$.

কারণ $p(x)^2$ সর্বদাই ≥ 0.

> Also,
>
> $$\begin{aligned} & \langle p(x), p(x)\rangle = 0 \\ \iff & \int_0^1 p(x)^2\,dx = 0 \\ \iff & p(x) \equiv 0 \quad [\because p \text{ is a polynomial}] \end{aligned}$$

এটা কেন হল, বুঝে নাও। ধরো এমন কোনো $a \in (0,1)$ আছে, যেখানে $p(a) \neq 0$. তবে $p(a)^2 > 0$. ধরো $p(a)^2$-কে ϵ নাম দিলাম। যেহেতু $p(x)^2$ একটা continuous function, তাই ওই a-কে ঘিরে এমন একটা neighbourhood পাবে, ধরো $(a-\delta, a+\delta) \subseteq [0,1]$, যার উপরে সর্বত্রই $p(x)^2 > \frac{\epsilon}{2}$ হবে[2]। সুতরাং--

$$\int_0^1 p(x)^2\,dx \geq \int_{a-\delta}^{a+\delta} p(x)^2\,dx \geq \int_{a-\delta}^{a+\delta} \tfrac{\epsilon}{2}\,dx \geq \epsilon\delta > 0.$$

কিন্তু বলে দিয়েছিল $\int_0^1 p(x)^2\,dx = 0$. তার মানে $p(x)$-টা $(0,1)$-এর উপরে সর্বত্রই শূন্য। এবার দেখাব যে, $p(x)$-টা পুরো $\mathbb{R}$-এর উপরেই সর্বত্র শূন্য, মানে $p(x)$ হল zero polynomial. যদি তা না হয়, তবে $p(x)$-এর একটা কোনো

[2]Real analysis শেখার সময়ে continuous function-দের sign-preserving property-র প্রমাণে যুক্তিটা দেখেছিলে।

degree আছে, যেটা একটা finite সংখ্যা, ধরো n. তবে আমরা জানি যে, তার সবচেয়ে বেশী n-খানা সংখ্যা থাকতে পারে, যেখানে $p(x)$ শূন্য হয়। কিন্তু এখানে তো দেখা যাচ্ছে যে $(0, 1)$-এর মধ্যে সব সংখ্যাতেই $p(x)$-টা শূন্য হয়ে বসে আছে। তার মানে $p(x)$-টা zero polynomial না হয়ে যায় না। ■

DAY 49
Inner product (part 3)

49.1 Orthogonal set

এই অধ্যায়ের গোড়ায় আমরা বলেছিলাম যে, একটা basis-এর vector-গুলো যদি পরস্পরের সঙ্গে orthogonal হয়, তবে তাদের নিয়ে কাজ করতে সুবিধা। যদি একগুচ্ছ vector পরস্পরের সঙ্গে orthogonal হয়, তবে সেই vector-দের set-টাকে বলে একটা **orthogonal set.**

DEFINITION: **Orthogonal set**

A subset A of an inner product space V is called an **orthogonal set** if

$$\forall \vec{u} \neq \vec{v} \in A \quad \langle \vec{u}, \vec{v} \rangle = 0.$$

যদি একটা orthogonal set-এর প্রতিটা vector-এর দৈর্ঘ্য (মানে norm) হয় 1, তবে তার একটা বিশেষ নাম আছে--

DEFINITION: **orthonormal set**

A subset A of an inner product space V is called orthonormal if

- $\forall \vec{u} \neq \vec{v} \in A \quad \langle \vec{u}, \vec{v} \rangle = 0.$
- $\forall \vec{v} \in A \quad \|\vec{v}\| = 1,$

যদি একটা basis আবার একটা orthogonal set-ও হয়, তবে সেরকম basis-কে আমরা বলি **orthogonal basis.** যদি $\mathbb{R}^2$, মানে গ্রাফ কাগজের, উদাহরণ দিয়ে ভাবো তবে orthogonal basis মানে হল দুটো axis পরস্পরের সঙ্গে perpendicular, কিন্তু দুই axis বরাবর একক দৈর্ঘ্য সমান নাও হতে পারে। দুটো উদাহরণ রয়েছে Fig 4-এ। এখানে গ্রাফ কাগজের ছকগুলো হবে rectangle-এর মত বা square-এর মত। যদি ছকগুলো সব square হয়, যাদের বাহুর দৈর্ঘ্য 1, তবে পাবে **orthonormal basis,** মানে এমন একটা orthogonal basis যেখানে প্রতিটা vector-এর norm হচ্ছে 1. যেমন Fig 4(b)-এর basis-টা একটা orthonormal basis হতে পারে। কিন্তু Fig 4(a)-তে basis-টা orthogonal হলেও, orthonormal নয়, কারণ দেখাই যাচ্ছে যে, vector দুটোর দৈর্ঘ্য আলাদা। আমরা ছোটোবেলায় সবসময়ে orthonormal basis-ই ব্যবহার করে এসেছি, কারণ কাজ করার পক্ষে সেটাই সবচেয়ে সুবিধাজনক।

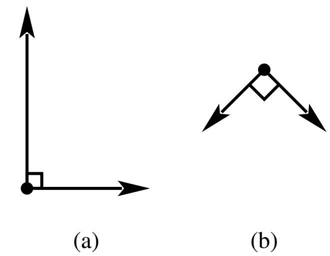


Fig 4

DEFINITION: Orthogonal/orthonormal basis

A basis of an inner product space is called an **orthogonal basis** if the vectors in it are mutually orthogonal. If the vectors also have norm 1, then the basis is called an **orthonormal basis** (**ONB**).

Example 10: Give an example of an orthogonal basis of $\mathbb{R}^4$.

SOLUTION: যে উদাহরণটা দেব তার চেয়ে সহজ আর কোনো ONB হয় না। এই ONB-টাকে বলে $\mathbb{R}^4$-এর **canonical basis** বা **Euclidean basis**.

One example of an orthogonal basis of $\mathbb{R}^4$ is

$$\{\ (1,0,0,0),\ \ (0,1,0,0),\ \ (0,0,1,0),\ \ (0,0,0,1)\ \}.$$

Here we are using the Euclidean inner product.

■

Example 11: Show that the set $\{(2,-3,1),\ (4,1,-5),\ (1,1,1)\}$ is an orthogonal set in the inner product space $\mathbb{R}^3$ with standard inner product.

SOLUTION: এখানে standard inner product বলতে Euclidean inner product বোঝানো হয়েছে।

Let

$$\begin{aligned}\vec{v}_1 &= (2,-3,1),\\ \vec{v}_2 &= (4,1,-5),\\ \vec{v}_3 &= (1,1,1).\end{aligned}$$

Then

$$\begin{aligned}\langle \vec{v}_1, \vec{v}_2 \rangle &= 2\times 4 + (-3)\times 1 + 1\times(-5) = 0,\\ \langle \vec{v}_1, \vec{v}_3 \rangle &= 2\times 1 + (-3)\times 1 + 1\times 1 = 0,\\ \langle \vec{v}_2, \vec{v}_3 \rangle &= 4\times 1 + 1\times 1 + (-5)\times 1 = 0.\end{aligned}$$

■

Exercise 13: Check if the set $\{(1,2,3),\ (-4,2,0),\ (1,1,1)\}$ is an orthogonal set in the inner product space $\mathbb{R}^3$ with standard inner product.
■

Exercise 14: Write down all the unit vectors in $\mathbb{R}^2$ that are orthogonal to $(3,4)$. ■

49.2 Independence-এর সাথে সম্পর্ক

Vector-দের তীর চিহ্ন দিয়ে ভাবলে দুটো vector-এর linearly independent হওয়া মানে তীর চিহ্ন দুটো একই লাইন বরাবর না হওয়া। Fig 5-এ দুটো vector দেখিয়েছি, যারা পরস্পরের সঙ্গে orthogonal. বলাই বাহুল্য যে, এরা এক লাইন বরাবর নেই। এ থেকে ফস্ করে মনে হতে পারে যে, কিছু vector পরস্পরের সঙ্গে orthogonal হলে linearly independent-ও হবে। এ কথাটা প্রায় ঠিক, খালি একটা সমস্যা আছে। সেটা এই যে, $\vec{0}$ সবার সঙ্গেই orthogonal, কিন্তু কোনো linearly independent set-এই $\vec{0}$ থাকতে পারে না। কেন $\vec{0}$ সবার সঙ্গে orthogonal, সেটা বোঝা খুব সহজ--

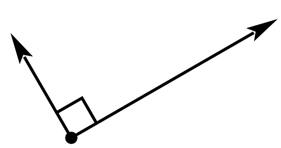
যেকোনো একটা $\vec{v}$ নাও। তবে $\left\langle \vec{0}, \vec{v} \right\rangle = \left\langle \vec{0}+\vec{0}, \vec{v} \right\rangle = \left\langle \vec{0}, \vec{v} \right\rangle + \left\langle \vec{0}, \vec{v} \right\rangle$. এবার দু দিক থেকে $\left\langle \vec{0}, \vec{v} \right\rangle$ কেটে দিলেই পাবে $\left\langle \vec{0}, \vec{v} \right\rangle = 0$.

কিন্তু যদি কিছু non-null vector পরস্পরের সঙ্গে orthogonal হয়, তবে আর চিন্তা নেই, ওরা linearly independent না হয়ে যায় না! সেটাই নীচের theorem-টার বক্তব্য।

> **THEOREM**
>
> Prove that an orthogonal set of non-null vectors in a Euclidean space V is linearly independent.

Proof:

Fig 5



Let $A \subseteq V$ be an orthogonal set of non-null vectors.

Shall show that A is linearly independent, by showing that for any distinct $\vec{v}_1, ..., \vec{v}_n \in A$ and any $a_1, ..., a_n \in F$ we have

$$\sum_i a_i \vec{v}_i = \vec{0} \implies \big(\forall k \in \{1, ..., n\} \quad a_k = 0\big).$$

Take any $k \in \{1, ..., n\}$.
Taking inner product with $\vec{v}_k$ of both sides of

$$\sum_i a_i \vec{v}_i = \vec{0},$$

we get

$$\left\langle \sum_i a_i \vec{v}_i, \vec{v}_k \right\rangle = \left\langle \vec{0}, \vec{v}_k \right\rangle,$$

or

$$\sum_i a_i \langle \vec{v}_i, \vec{v}_k \rangle = 0,$$

or

$$a_k \langle \vec{v}_k, \vec{v}_k \rangle = 0,$$

Since $\forall i \neq k$ we have $\langle \vec{v}_i, \vec{v}_k \rangle = 0$.
Now $\because \vec{v}_k \neq \vec{0}, \therefore \langle \vec{v}_k, \vec{v}_k \rangle \neq 0$,
$\therefore a_k = 0$.
Since $k \in \{1, ..., n\}$ was arbitrary, we have $a_1 = \cdots = a_n = 0$, as required.

[Q.E.D]

Example 12: আমরা একটু আগেই একটা অংকে দেখলাম যে, $\{(2, -3, 1), \ (4, 1, -5), \ (1, 1, 1)\}$ হল একটা orthogonal set. এখানে আমরা $\mathbb{R}^3$-এর উপরে Euclidean inner product নিয়ে কাজ করছি। পরীক্ষা করে দেখো এই set-টা $\mathbb{R}^3$-র একটা basis হচ্ছে কিনা। **(2012.6c)**

SOLUTION: আমরা জানি যে, nonzero vector-রা orthogonal হলে linearly independent-ও হয়। সুতরাং এই set-টা linearly independent. এর সাইজ হল 3. আবার $\mathbb{R}^3$-র dimension-ও 3. সুতরাং set-টা অবশ্যই $\mathbb{R}^3$-এর একটা basis হবে। ■

Exercise 15: $S = \{\alpha_1, ..., \alpha_n\}$ is a set of non-null orthogonal vectors of $\mathbb{R}^n$. Show that S is a basis of $\mathbb{R}^n$.

HINT:
আগের অংকের মতই। ■

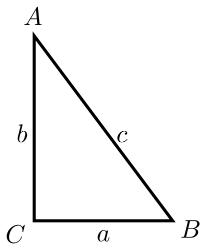


Fig 6

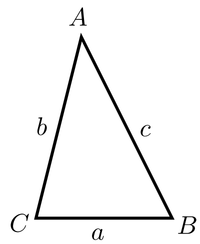


Fig 7

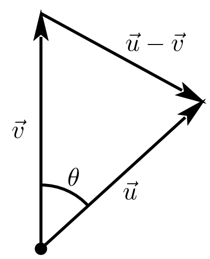


Fig 8

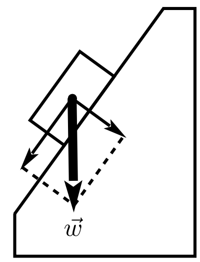


Fig 9

49.3 Inner product-এর জ্যামিতিক ব্যাখ্যা

দুটো vector-এর inner product-এর একটা জ্যামিতিক ব্যাখ্যা দেওয়া যায়, ওদের মধ্যেকার কোণ ব্যবহার করে। এবার সেটা শিখব। তবে ব্যাখ্যাটা খালি $\mathbb{R}^2$ বা $\mathbb{R}^3$-এর বেলাতেই খাটে, কারণ এরাই হল আমাদের পরিচিত জ্যামিতির জগৎ।
শুরু করব এমন একটা জিনিস নিয়ে, যেটা আমরা হায়ার সেকণ্ডারীতে শিখি, cosine rule. এটা বস্তুতঃ Pythagoras' theorem-এরই মত একটা জিনিস, খালি আরও general, যেকোনো triangle-এর জন্যই খাটে, right angled না হলেও। Pythagoras' theorem-টা ছিল (Fig 6)

$$c^2 = a^2 + b^2.$$

যদি C কোণটা 90° না হয়, তবে তার জন্য খালি একটা বাড়তি জিনিস লাগবে (Fig 7)--

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

এটাই হল cosine rule. এর অন্যতম প্রয়োগ হল একটা ত্রিভুজের বাহুদের দৈর্ঘ্য জানা থাকলে তা থেকে কোণ বার করা, কারণ cosine rule-টাকে এভাবে লেখা যায়--

$$\cos C = \frac{a^2 + b^2 - c^2}{2ab}.$$

এখানে বাঁদিকে খালি কোণ আছে, আর ডানদিকে রয়েছে খালি বাহু তিনটের দৈর্ঘ্য। এইটাই এবার আমরা প্রয়োগ করব। ধরো $\vec{u}$ আর $\vec{v}$ দুটো nonzero vector. (দুজনকেই nonzero নিয়েছি, নইলে ওদের মধ্যে কোণ মাপা যাবে না)। ফলে আমরা একটা ত্রিভুজ পাচ্ছি (Fig 8), যার বাহুর দৈর্ঘ্যগুলো হল $\|\vec{u}\|, \|\vec{v}\|$ আর $\|\vec{u}-\vec{v}\|$. এখানে আমরা Euclidean inner product নিয়ে কাজ করছি। সুতরাং

$$\cos\theta = \frac{\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2}{2\|\vec{u}\|\cdot\|\vec{v}\|} = \frac{\langle \vec{u}, \vec{v}\rangle}{\|\vec{u}\|\cdot\|\vec{v}\|},$$

কারণ

$$\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2 = \cdots = 2\langle \vec{u}, \vec{v}\rangle.$$

এখান থেকে এটাও বুঝতে পারবে, কেন $\theta = 90^\circ$ হলে $\langle \vec{u}, \vec{v}\rangle = 0$ হয়, কারণ $\cos 90^\circ = 0$. তবে আরেক বার মনে করিয়ে দিই যে, কোণের ধারণাটা খালি $\mathbb{R}^2$ আর $\mathbb{R}^3$-এর বেলাতেই খাটে, যদিও inner product আর norm-এর ধারণাটা অনেক বেশী general.

49.4 Component বার করা

হায়ার সেকণ্ডারীর physics-এ অনেক সময়ে একটা vector দেওয়া থাকত, এবং কোনো একটা দিক বরাবর তার component (উপাংশ) বার করতে হত। যদি মনে না পড়ে, তবে Fig 9 দেখো। এখানে একটা inclined plane (আনত তল)-এর উপরে একটা বাক্স রাখা আছে। আমরা বাক্সের ওজন $\vec{w}$-কে দুটো component-এ ভেঙে নিয়েছি, একটা component আছে inclined plane বরাবর, অন্যটা তার সাথে perpendicular (বা আমরা এখন যেটাকে orthogonal বলছি)। Component বার করা বলতে ঠিক এটাই বোঝাচ্ছি, Fig 10 দেখো--

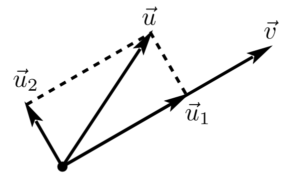


Fig 10

তোমাকে দুটো vector দেওয়া থাকবে, $\vec{u}$ আর $\vec{v}$, (এর মধ্যে $\vec{v} \neq \vec{0}$). তোমাকে বলা হবে $\vec{v}$-এর লাইন বরাবর $\vec{u}$-এর component বার করতে। এর মানে হল $\vec{u}$-কে এমন দুইভাগে ভাঙতে হবে--

$$\vec{u} = \vec{u}_1 + \vec{u}_2 \tag{*}$$

যাতে $\vec{u}_1$ থাকে $\vec{v}$-এর লাইন বরাবর (মানে $\vec{u}_1$ হয় $\vec{v}$-এর multiple), আর $\vec{u}_2$ হয় $\vec{v}$-এর সঙ্গে orthogonal.

এই কাজটা সব সময়েই করা যাবে, এবং কায়দাটা আমরা একটু আগে ব্যবহারও করেছি। কায়দাটা এরকম--প্রথমে (*)-এর দুই পাশেই $\vec{v}$-এর সঙ্গে inner product নাও--

$$\langle \vec{u}, \vec{v} \rangle = \langle \vec{u}_1 + \vec{u}_2, \vec{v} \rangle = \langle \vec{u}_1, \vec{v} \rangle + \langle \vec{u}_2, \vec{v} \rangle .$$

এদের মধ্যে ডানদিকের দ্বিতীয় inner product-টা হল 0, কারণ $\vec{u}_2$-কে $\vec{v}$-এর সঙ্গে orthogonal হতে হবে, অতএব--

$$\langle \vec{u}, \vec{v} \rangle = \langle \vec{u}_1, \vec{v} \rangle .$$

এদিকে আমরা জানি যে $\vec{u}_1$ হল $\vec{v}$-এর কোনো multiple, ধরো $\vec{u}_1 = \lambda \vec{v}$, তাহলে $\langle \vec{u}_1, \vec{v} \rangle = \langle \lambda \vec{v}, \vec{v} \rangle = \lambda \langle \vec{v}, \vec{v} \rangle$. তাই

$$\lambda = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}.$$

এখানে আমরা $\langle \vec{v}, \vec{v} \rangle$ দিয়ে নিশ্চিন্তে ভাগ করতে পারলাম, কারণ $\vec{v} \neq \vec{0}$. সুতরাং $\vec{u}_1 = \lambda \vec{v}$ এবার বার করে ফেলা যাবে অনায়াসেই--

$$\vec{u}_1 = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v}.$$

Exercise 16: সত্যিই এবার হাতেকলমে পরীক্ষা করে দেখো যে, $\vec{u}_2 = \vec{u} - \vec{u}_1$-টা $\vec{v}$-এর সঙ্গে orthogonal হচ্ছে। ■

এই জিনিসটা physics, vector analysis থেকে শুরু করে computer graphics ইত্যাদি বিভিন্ন জায়গায় খুব কাজে লাগে। তাই একটা theorem হিসেবে লিখে রাখি--

THEOREM

Let V be an inner product space. Let $\vec{u}, \vec{v} \in V$ where $\vec{v} \neq \vec{0}$. Then the component of $\vec{u}$ along $\vec{v}$ is

$$\frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v}.$$

এখানে একটা ব্যাপারে সাবধান। আমরা $\langle\vec{u},\vec{v}\rangle$-এর মধ্যে $\vec{v}$-কে ডানদিকে রেখেছি। যদি field-টা $\mathbb{R}$ হয়, তবে বাঁদিক-ডানদিকে কিছু এসে যায় না। কিন্তু field-টা $\mathbb{C}$ হলে এটা গুরুত্বপূর্ণ--যে vector-টা বরাবর component নিচ্ছ, সেটা যেন inner product নেবার সময়ে ডানদিকে থাকে।

49.5 Quotient space-এর প্রয়োগ

আমরা এ বইয়ের গোড়ার দিকে quotient space-এর কথা শিখেছিলাম। এবার তার একটা প্রয়োগ আলোচনা করব। এই অংশটুকু বইয়ের বাকি সব কিছুর তুলনায় বেশ অনেকটাই কঠিন, এবং এ বইতে পরে কোথাও লাগবেও না। সুতরাং চাইলে আজকের অবশিষ্ট পড়াটুকু বাদ দিয়ে যেতে পারো।

আমরা $\mathbb{R}^n$-এর উপরে Euclidean norm-এর কথা জানি, $\|\vec{x}\| = \sqrt{\sum_i x_i^2}$, যেখানে $\vec{x} = (x_1, ..., x_n)$. যদি আমরা x_i-এর বদলে $x(i)$ লিখি, তবে লক্ষ্য করো, $\vec{x}$ আসলে একটা function, যার domain হল $\{1, ..., n\}$ আর codomain হল $\mathbb{R}$. মানে $\vec{x} : \{1, ..., n\} \to \mathbb{R}$. গণিতজ্ঞরা অনেক সময়ে $\{1, ..., n\}$-এর বদলে অন্য কোনো set নিয়েও একইভাবে norm বানাবার চেষ্টা করেন। যেমন ধরো যদি $[0, 1]$ নিই, তবে আমাদের vector-রা হবে $f : [0, 1] \to \mathbb{R}$ জাতীয় function. $\mathbb{R}^n$-এর অনুকরণে এদের norm-এর সংজ্ঞা আমরা কল্পনা করতে পারি এইরকম--

$$\|f\| \stackrel{?}{=} \sqrt{\int_0^1 f(x)^2\, dx}. \qquad (*)$$

ওই প্রশ্নচিহ্নটার মানে হল, আমি বলছি না যে, এটা norm হবেই, এটা খালি অনুকরণের প্রথম প্রচেষ্টামাত্র। এই প্রথম প্রচেষ্টা যে ধোপে টিঁকবে না, সেটা একটা জিনিস ভাবলেই বোঝা যায়-- কারণ $f(x)^2$ যে $[0, 1]$-এর উপরে integrable হবে, এমন কোনো গ্যারান্টি নেই। ঠিক আছে, সেই ত্রুটি সংশোধন করার জন্য আমরা না হয় একটা বাড়তি শর্ত চাপালাম যে, আমরা কেবলমাত্র সেই সব $f : [0, 1] \in \mathbb{R}$ নিয়েই কাজ করব, যাদের জন্য $f(x)^2$ হবে integrable. যেসব function এই শর্ত পালন করে, তাদের set-কে বলে $\mathscr{L}_2[0, 1]$. ওই অক্ষরটা হল L-এর একটা বাহারি সংস্করণ। এটা দেখানো যায় যে, $\mathscr{L}_2[0, 1]$ একটা vector space. এবার তবে দ্বিতীয় প্রচেষ্টা--আমরা কি (*) ব্যবহার করে $\mathscr{L}_2[0, 1]$-এর উপরে একটা norm বানাতে পারি?

দুঃখের ব্যাপার, একটা ছোট্টো সমস্যা হবে। যদি কোনো vector-এর norm হয় 0, তবে সেই vector-টাকে $\vec{0}$ হতে হবে। কিন্তু এখানে যদি আমি $f : [0, 1] \to \mathbb{R}$ এইভাবে নিই--

$$f(x) = \begin{cases} 1 & \text{if } x = \frac{1}{2} \\ 0 & \text{otherwise} \end{cases},$$

তবে $\int_0^1 f(x)^2\, dx = 0$ হচ্ছে, কিন্তু $f(x) \not\equiv 0$. এই সমস্যা থেকে রেহাই পেতে আমরা একটা কৌশল করি, এবং সেখানেই quotient space লাগে। আমরা প্রথমে বিশেষ করে সেই সব $f \in \mathscr{L}_2[0, 1]$-গুলো নেব, যাদের জন্য $\int_0^1 f(x)^2\, dx = 0$ হয়। একটু ভাবলেই বুঝবে যে, এদের set-টা $\mathscr{L}_2[0, 1]$-এর একটা subspace হচ্ছে। ধরো এর নাম দিলাম S. তবে আমরা $\mathscr{L}_2[0, 1]/S$ নিয়ে কাজ করব। এটাও একটা vector space. এইবার এটা সহজেই দেখানো যায় যে, (*) ব্যবহার করলে এই vector space-এর উপরে সত্যিই একটা norm পাওয়া যায়।

আমরা এখানে Riemann integration নিয়ে কাজ করছিলাম। লোকে আজকাল অধিকাংশ ক্ষেত্রেই Lebesgue integration নামে আরো general এক ধরণের integration নিয়ে কাজ করে। সেখানেও একই জিনিস করা যায়, এবং সেক্ষেত্রে এই $\mathscr{L}_2[0, 1]/S$-র একটা নাম আছে। একে বলে $L_2[0, 1]$. অংকের দুনিয়ায় এর ভারী কদর।

এই আলোচনাটা যদি তোমার জটিল লেগে থাকে, তবে ঘাবড়ে যেও না। এই ধারণাটা এ বইতে পরে আর কোথাও লাগবে না।

DAY 50
Inner product (part 4)

এবার আমরা দুটো গুরুত্বপূর্ণ inequality শিখব--Cauchy-Schwartz inequality আর triangle inequality.

50.1 Cauchy-Schwartz inequality

একটু আগে inner product-এর জ্যামিতিক ব্যাখ্যা শেখার সময়ে আমরা দেখলাম যে, $\mathbb{R}^2$ বা $\mathbb{R}^3$-র ক্ষেত্রে $\frac{\langle \vec{u}, \vec{v}\rangle}{\|\vec{u}\|\cdot\|\vec{v}\|}$-কে আমরা $\cos\theta$ বলে ভাবতে পারি, যেখানে θ হল $\vec{u}$ আর $\vec{v}$-এর মধ্যেকার কোণ। আমরা জানি যে θ যাই হোক না কেন, সর্বদাই $|\cos\theta| \leq 1$ হতে বাধ্য। সুতরাং

$$\left|\frac{\langle \vec{u}, \vec{v}\rangle}{\|\vec{u}\|\cdot\|\vec{v}\|}\right| \leq 1$$

হবে, মানে square নিলে

$$\langle \vec{u}, \vec{v}\rangle^2 \leq \|\vec{u}\|^2 \cdot \|\vec{v}\|^2 \qquad (*)$$

হবে। আমরা জানি যে $\cos\theta$-র উপমাটা $\mathbb{R}^2$ বা $\mathbb{R}^3$-এর বাইরে খাটে না। মজার কথা হল (*) কিন্তু যেকোনো inner product space-এই খাটে। একেই বলে Cauchy-Schwartz inequality.

Cauchy-Schwartz inequality

Let V be any inner product space. Let $\vec{u}, \vec{v} \in V$. Then

$$|\langle \vec{u}, \vec{v}\rangle| \leq \|\vec{u}\|\cdot\|\vec{v}\|,$$

and equality holds if and only if $\{\vec{u}, \vec{v}\}$ is linearly dependent.

Proof: প্রমাণটা করা আগে একটা সহজ উপমা ভেবে নিই। ধরো তোমায় Fig 11 দিয়ে বলা হল $\cos\theta \leq 1$ দেখাতে। কী করে করবে? একটা কায়দা হল সরাসরি সংজ্ঞা লাগানো--

$$\cos\theta = \frac{BC}{AC}.$$

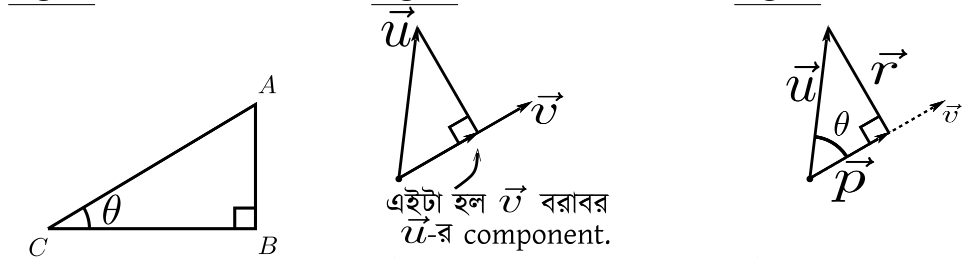
যেহেতু AC হল hypotenuse (অতিভুজ) তাই অবশ্যই $BC \leq AC$, সুতরাং $\cos\theta \leq 1$. আর BC যে সত্যিই AC-র চেয়ে ছোটো সেটার প্রমাণ চাইলে আমরা Pythagoras লাগাতে পারি--

$$AC = \sqrt{BC^2 + AB^2} \geq BC,$$

যেহেতু AB^2 তো আর < 0 হতে পারে না। এ থেকে আরেকটা জিনিসও চলে আসে, $\cos\theta = 1$ হবে, যখন $AB = 0$ হবে, মানে BC আর AC পরস্পরের সঙ্গে মিশে যাবে। এই উপমাটা মাথায় রাখলেই Cauchy-Schwartz inequality-র প্রমাণটা অনায়াসে বুঝতে পারবে।

আমরা এখানে একটা right angled triangle নিয়ে কাজ করব। Fig 12 দেখো। এর একটা বাহু হল $\vec{u}$, আরেকটা $\vec{v}$ বরাবর $\vec{u}$-র component. অবশ্য $\vec{v}$ বরাবর component বার করা তখনই সম্ভব, যখন $\vec{v} \neq \vec{0}$ হবে। সুতরাং $\vec{v} = \vec{0}$ কেসটাকে আগে আলাদা করে ভেবে নিই--

Fig 11 **Fig 12** **Fig 13**



> The result holds trivially if $\|\vec{v}\| = 0$ (ie, if $\vec{v} = \vec{0}$), because then both sides are 0.

এবার $\vec{v} \neq \vec{0}$ হলে কী হয় দেখি--

> So let us assume $\vec{v} \neq \vec{0}$.

এবার আমরা Fig 12 নিয়ে কাজ করব। একটা বাহু তো $\vec{u}$ আছেই, অন্য দুটোরও নাম দিয়ে নিই--

> Let
> $$\vec{p} = \underbrace{\frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}}_{a} \vec{v} \text{ and } \vec{r} = \vec{u} - \vec{p}.$$

Fig 13 থেকেই দেখা যাচ্ছে যে, $\vec{p}$ হল $\vec{v}$ বরাবর $\vec{u}$-এর component, এবং $\vec{p}$ আর $\vec{r}$ পরস্পরের সঙ্গে orthogonal. সেটা অংক কষে দেখিয়ে রাখি--

> Then
> $$\langle \vec{r}, \vec{p} \rangle = \langle \vec{u} - a\vec{v}, a\vec{v} \rangle = \bar{a} \langle \vec{u}, \vec{v} \rangle - a\bar{a} \langle \vec{v}, \vec{v} \rangle = \bar{a} \big(\underbrace{\langle \vec{u}, \vec{v} \rangle - a \langle \vec{v}, \vec{v} \rangle}_{0} \big) = 0.$$

এবার Pythagoras-এর কায়দাটা আসছে--

> Now $\|\vec{r}\|^2 \geq 0$. So
> $$\begin{aligned} & \langle \vec{u} - a\vec{v}, \vec{u} - a\vec{v} \rangle \geq 0 \\ \text{or,} \quad & \langle \vec{u}, \vec{u} \rangle - \langle \vec{u}, a\vec{v} \rangle - \langle a\vec{v}, \vec{u} \rangle + \langle a\vec{v}, a\vec{v} \rangle \geq 0 \\ \text{or,} \quad & \|\vec{u}\|^2 - 2\mathrm{Re}\big(a \langle \vec{v}, \vec{u} \rangle \big) + |a|^2 \|\vec{v}\|^2 \geq 0. \end{aligned}$$

এবার a-ওয়ালা জিনিস দুটো ঠিক কী, বোঝা যাক। আমরা জানি $a = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}$. সুতরাং--

> Now
> $$|a|^2 = \left| \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \right|^2 = \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\langle \vec{v}, \vec{v} \rangle^2} = \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\|\vec{v}\|^4},$$
> since $\langle \vec{v}, \vec{v} \rangle > 0$.
> So $|a|^2 \|\vec{v}\|^2 = \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\|\vec{v}\|^2}$.
> Also, $a \langle \vec{v}, \vec{u} \rangle = \frac{\langle \vec{u}, \vec{v} \rangle \langle \vec{v}, \vec{u} \rangle}{\|\vec{v}\|^2} = \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\|\vec{v}\|^2}$.

এটা একটা real সংখ্যা, তাই--

Hence

$$\mathrm{Re}\big(a\,\langle\vec{v},\vec{u}\rangle\,\big)=\frac{|\,\langle\vec{u},\vec{v}\rangle\,|^2}{\|\vec{v}\|^2}.$$

আমাদের হাতে ছিল $\|\vec{u}\|^2-2\mathrm{Re}\big(a\,\langle\vec{v},\vec{u}\rangle\,\big)+|a|^2\|\vec{v}\|^2\geq 0$. দেখা যাক সেটা এখন কীসে পরিণত হয়।

Thus,

$$\|\vec{u}\|^2-2\frac{|\,\langle\vec{u},\vec{v}\rangle\,|^2}{\|\vec{v}\|^2}+\frac{|\,\langle\vec{u},\vec{v}\rangle\,|^2}{\|\vec{v}\|^2}\geq 0$$

$$\text{or,}\quad \|\vec{u}\|^2-\frac{|\,\langle\vec{u},\vec{v}\rangle\,|^2}{\|\vec{v}\|^2}\geq 0$$

$$\text{or,}\quad \|\vec{u}\|^2\geq\frac{|\,\langle\vec{u},\vec{v}\rangle\,|^2}{\|\vec{v}\|^2}$$

$$\text{or,}\quad |\,\langle\vec{u},\vec{v}\rangle\,|\leq\|\vec{u}\|\,\|\vec{v}\|\qquad[\;\because\|\vec{u}\|\,,\|\vec{v}\|\geq 0]\,,$$

as required.

এবার কখন equality হয়, দেখব। তার জন্য সেই right-angled triangle-টার ছবি মাথায় রাখতে হবে। আমরা চাইছি যেন Fig 13-এ $\cos\theta=1$ হয়। তার মানে $\vec{r}=\vec{0}$ হয়ে গিয়ে বাকি দুটো বাহু দুটো পরস্পরের সঙ্গে একেবারে মিশে এক হয়ে যায়, অর্থাৎ $\vec{u}$ আর $\vec{v}$ এক লাইন বরাবর হয়। এই ছবিটা যে আঁকতে পারছি, তার কারণ আমরা $\vec{v}\neq\vec{0}$ নিয়ে কাজ করছি ($\vec{v}=\vec{0}$ কেসটা গোড়াতেই দেখে নিয়েছিলাম)।

Equality condition:

Let equality hold. Then $\|\vec{r}\|^2=0$,

So $\vec{u}=\vec{p}$, ie, $\vec{u}=a\vec{v}$.

Hence $\{\vec{u},\vec{v}\}$ is linearly independent.

Conversely, if $\{\vec{u},\vec{v}\}$ is linearly independent, then $\vec{u}=b\vec{v}$ for some scalar b ($\because\vec{v}\neq 0$).

Then $\langle\vec{u},\vec{v}\rangle=\langle b\vec{v},\vec{v}\rangle=b\,\|\vec{v}\|^2$.

Also, $\|\vec{u}\|=|b|\cdot\|\vec{v}\|$.

So $|\,\langle\vec{u},\vec{v}\rangle\,|=|b|\cdot\|\vec{v}\|^2=|b|\cdot\|\vec{v}\|\cdot\|\vec{v}\|=\|\vec{u}\|\cdot\|\vec{v}\|$, as required.

[Q.E.D]

50.2 Triangle inequality

Fig 14-এ একটা ঘড়ি দেখিয়েছি। মিনিটের কাঁটার দৈর্ঘ্য 5 cm, ঘন্টার কাঁটার দৈর্ঘ্য 3 cm. আমরা জানতে চাই কাঁটা দুটোর ডগার মধ্যে দূরত্বটা (ড্যাশ্ ড্যাশ্ লাইনটা) সবচেয়ে বেশী আর সবচেয়ে কম কত হতে পারে। নিশ্চয়ই সন্দেহ নেই যে, দূরত্বটা সর্বাধিক হবে যখন কাঁটা দুটো Fig 15-এর মত থাকবে। তখন দূরত্বটা হচ্ছে কাঁটাদুটোর দৈর্ঘ্যের যোগফল। আবার সবচেয়ে কম হবে Fig 16-এর মত অবস্থায়। সেখানে দূরত্বটা হবে দৈর্ঘ্য দুটোর পার্থক্য। যদি কাঁটা দুটোর জায়গায় দুটো vector নিই, $\vec{u}$ আর $\vec{v}$, তবে ওদের দৈর্ঘ্য হবে $\|\vec{u}\|$ আর $\|\vec{v}\|$, আর ওদের নিজেদের মধ্যে দূরত্ব হবে $\|\vec{u}-\vec{v}\|$. সুতরাং আমরা পাচ্ছি

$$\big|\;\|\vec{u}\|-\|\vec{v}\|\;\big|\leq\|\vec{u}-\vec{v}\|\leq\|\vec{u}\|+\|\vec{v}\|.$$

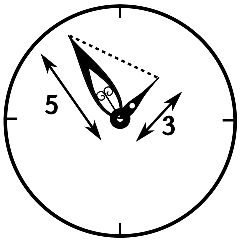


Fig 14

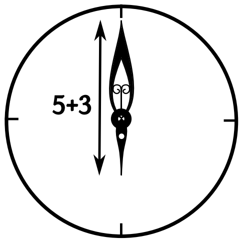


Fig 15

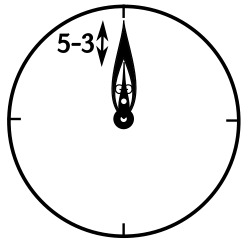


Fig 16

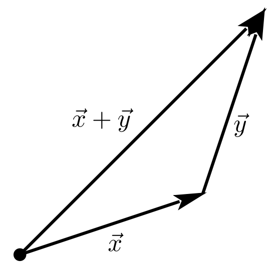


Fig 17

এটাকেই বলে triangle inequality. এখানে প্রথমে ওই $\big|\ \|\vec{u}\| - \|\vec{v}\|\ \big|$-টা দেখে ঘাবড়ে যেও না, ওটার মানে হল দুটো কাঁটার দৈর্ঘ্যের পার্থক্য। যেহেতু বুঝতে পারছি না $\|\vec{u}\|$ আর $\|\vec{v}\|$-এর মধ্যে কে বড়, তাই $\|\vec{u}\| - \|\vec{v}\|$-এর absolute value নিয়ে নিয়েছি।

Triangle inequality অংকের দুনিয়ায় অত্যন্ত গুরুত্বপূর্ণ জিনিস। শুধু linear algebra-তেই নয়, real analysis, metric space ইত্যাদি অংকের অন্যান্য শাখাতেও এর প্রচুর প্রয়োগ। সব প্রয়োগেরই মূল কথা একই। দুটো জিনিসের মধ্যে দূরত্ব কত হবে সেটার একটা আন্দাজ পাওয়া, যেখানে জিনিস দুটোর প্রত্যেকেই তৃতীয় কোনো জিনিস থেকে কত দূরে আছে, সেটুকুমাত্র বলা আছে। আমাদের ঘড়ির উদাহরণে যেমন কেন্দ্রটা ছিল এই তৃতীয় জিনিস। Vector-এর বেলায় $\vec{0}$ সেই ভূমিকা পালন করছিল (মনে রেখো $\|\vec{u}\|$ হল $\vec{0}$ থেকে $\vec{u}$-এর দূরত্ব)।

যেভাবে আমরা triangle inequality-টা লিখেছি, তাতে এর পিছনের গল্পটা স্পষ্ট হলেও দেখতে একটু বিচ্ছিরি। একই জিনিসকে একটু সুন্দর করে লেখা যায় এইভাবে--

Triangle inequality

Let V be any inner product space, and let $\vec{x}, \vec{y} \in X$. Then

$$\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|.$$

Equality holds if and only if either $\vec{x} = \lambda\vec{y}$ or $\vec{y} = \lambda\vec{x}$ for some real number $\lambda \geq 0$.

Triangle inequality-র এই রূপটাকেও ছবি দিয়ে বোঝা বেশ সহজ (Fig 17)। $\vec{x}, \vec{y}$ আর $\vec{x}+\vec{y}$ এই তিনটে vector-কে নিয়ে একটা ত্রিভুজ কল্পনা কর। তাহলে একটা বাহুর দৈর্ঘ্য হল $\|\vec{x}+\vec{y}\|$, সেটা অন্যদুটো বাহুর মোট দৈর্ঘ্যকে ছাড়িয়ে যেতে পারে না। আর সমান হবে, যদি $\vec{x}$ এবং $\vec{y}$ দুজনেই একই সরলরেখা বরাবর একই দিকে মুখ করে থাকে।

Proof:

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \langle x,x\rangle + 2\mathrm{Re}(\langle x,y\rangle) + \langle y,y\rangle \quad \left[\because \langle x,y\rangle = \overline{\langle y,x\rangle}\right] \\
&\leq \langle x,x\rangle + 2|\langle x,y\rangle| + \langle y,y\rangle \quad \left[\because \forall a \in \mathbb{C} \quad \mathrm{Re}(a) \leq |a|\right] \\
&\leq \langle x,x\rangle + 2\|x\|\|y\| + \langle y,y\rangle \quad [\text{by Cauchy-Schwartz inequality}] \\
&= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= \big(\|x\| + \|y\|\big)^2.
\end{aligned}$$

> Hence
> $$\|x+y\| \le \|x\| + \|y\|,$$
> as required, because norms are always nonnegative.

এবার equality-র খবর নিতে হবে। এখানে inequality এসেছে দুই ধাপে--

$$\mathrm{Re}(\langle x, y\rangle) \le |\langle x, y\rangle| \le \|\vec{x}\| \cdot \|\vec{y}\|.$$

সুতরাং triangle inequality-তে equality হবে if and only if এই দুই জায়গাতেই equality হয়।

> Condition for equality: Equality holds if and only if
> $$\mathrm{Re}(\langle x, y\rangle) = |\langle x, y\rangle| \text{ and } |\langle x, y\rangle| = \|\vec{x}\| \cdot \|\vec{y}\|.$$

প্রথম inequality-টা equality হবে, যখন $\langle \vec{x}, \vec{y}\rangle$-এর real অংশ আর modulus সমান হবে। কোন্ complex সংখ্যাদের বেলায় সেটা হয়? উত্তর হল--যখন সংখ্যাটা আসলে একটা non-negative, real সংখ্যা হবে। সুতরাং এখানে আমাদের দরকার--

> The first equality holds if and only if $\langle \vec{x}, \vec{y}\rangle \ge 0$, since for $z \in \mathbb{C}$ we have $\mathrm{Re}(z) = |z| \iff z$ is nonneg real.

যদি আমরা $\langle \vec{x}, \vec{y}\rangle$-কে $\cos\theta$ দিয়ে ভাবি, তবে $\langle \vec{x}, \vec{y}\rangle \ge 0$ মানে $\vec{x}, \vec{y}$-এর মধ্যে কোণটা হতে হবে একটা acute angle (সূক্ষ্ম কোণ)। এবার দ্বিতীয় inequality-টা দেখি। এটা তো Cauchy-Schwartz inequality. আমরা জানি এখানে equality হওয়া মানে $\vec{x}$ আর $\vec{y}$ একই লাইন বরাবর থাকবে।

> By the Cauchy-Schwartz inequality, the second equality holds if and only if $\{\vec{x}, \vec{y}\}$ is linearly dependent.

দুটো vector যদি একই লাইনে থাকে, তবে ওদের মধ্যেকার কোণ হবে $0°$ বা $180°$. এদিকে কোণটাকে acute-ও হতে হবে। তার মানে কোণটাকে $0°$-ই হতে হবে, অর্থাৎ $\vec{x}$ আর $\vec{y}$-কে একই দিকে মুখ করে থাকতে হবে।

> Combining, equality holds in triangle inequality if and only if $\vec{x}$ and $\vec{y}$ are both along the same line and in the same direction.

[Q.E.D]

50.3 Orthogonal/unitary matrix

আমরা প্রথম অধ্যায়েই orthogonal আর unitary matrix-দের কথা জেনেছি। যদি A এমন একটা real matrix হয়, যাতে $A'A = AA' = I$ হয়, তবে সেরকম A-কে বলে orthogonal matrix. আবার যদি A এমন একটা complex matrix হয়, যাতে $A^*A = AA^* = I$ হচ্ছে, তবে তাকে বলব একটা unitary matrix. বুঝতেই পারছ যে, $A'A = AA'$ বা $A^*A = AA^* = I$ হবার জন্য A-কে square matrix হতেই হবে। এখানে "orthogonal" শব্দটার ব্যবহার শুনেই আন্দাজ করতে পারছ যে, এর সঙ্গে inner product-এর কিছু একটা সম্পর্ক আছে। যদি A-র সাইজ $n \times n$ হয়, তবে আমরা orthogonal matrix-এর বেলায় $\mathbb{R}^n$-এর উপর Euclidean inner product নিয়ে কাজ করব। আর unitary-র বেলায় কাজ করব $\mathbb{C}^n$-এর উপর Euclidean inner product-এর complex সংস্করণ নিয়ে, মানে $\langle \vec{x}, \vec{y}\rangle = \vec{y}^*\vec{x}$ নিয়ে।

এবার সম্পর্কটা বলি--যদি A একটা orthogonal matrix হয় তবে ওর row-গুলো মিলে একটা ONB হবে, মানে প্রত্যেকটার row-এর norm হবে 1, এবং ওরা পরস্পরের সঙ্গে orthogonal হবে। একই ভাবে column-রা মিলেও একটা ONB হবে। বিপরীতপক্ষে যদি row-গুলো মিলে (বা column-গুলো মিলে) একটা ONB হয়, তবে matrix-টা orthogonal হতে বাধ্য।
তার মানে একটা $n \times n$ সাইজের real matrix A-র জন্য এই তিনটে শর্ত আসলে সমার্থক--

1. A-টা orthogonal matrix,
2. A-র row-গুলো মিলে $\mathbb{R}^n$-এর একটা ONB হয়।
3. A-র column-গুলো মিলে $\mathbb{R}^n$-এর একটা ONB হয়।

এদের মধ্যে যেকোনো একটা শর্ত খাটলেই বাকি দুটোও খাটতে বাধ্য। একইরকম কথা খাটবে unitary matrix-দের বেলাতেও, খালি সেখানে কাজ হবে $\mathbb{C}^n$-এর মধ্যে।
এটা প্রমাণ করা খুবই সহজ, কিন্তু প্রমাণের আগে একটা উদাহরণ দেখে নিই।

Example 13: প্রথমে দেখাও যে $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ একটা orthogonal matrix. এবার এর row-দুটোকে তীরচিহ্ন হিসেবে আঁকো। ওদের norm, মানে দৈর্ঘ্য কী 1 হচ্ছে? ওরা কি পরস্পরের সঙ্গে $90°$-তে আছে (মানে orthogonal)? আর column-এর বেলায়?

SOLUTION: AA' আর $A'A$ বার করলেই দেখবে যে দুক্ষেত্রেই I হচ্ছে। সুতরাং orthogonal. বস্তুতঃ যেহেতু A একটা square matrix, তাই AA' আর $A'A$-র মধ্যে যেকোনো একটাকে I দেখালেই অন্যটাও আপনা থেকেই I হতে বাধ্য। যেমন ধরো AA' হল

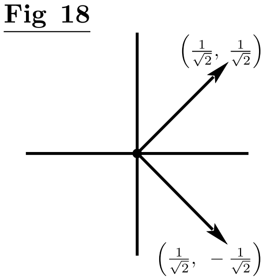
$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} + (-\frac{1}{\sqrt{2}}) \times (-\frac{1}{\sqrt{2}}) = 1 & \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} + (-\frac{1}{\sqrt{2}}) \times \frac{1}{\sqrt{2}} = 0 \\ \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \times (-\frac{1}{\sqrt{2}}) = 0 & \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} = 1 \end{bmatrix}.$$

এবার row দুটোকে তীর চিহ্ন হিসেবে আঁকব। Row দুটো হল $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ আর $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. ছবি আঁকলে Fig 18-এর মত। ওদের মাঝখানের কোণটা সত্যিই $90°$, আর দুজনেরই দৈর্ঘ্য 1. এইটা অবশ্য ছবি ছাড়া সরাসরিও inner product নিয়ে দেখানো যেত। যেমন ধরো, দুটো row-এর মধ্যে inner product হল

$$\left\langle \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \right\rangle = \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \times \left(-\frac{1}{\sqrt{2}}\right) = 0. \qquad (*)$$

Fig 18



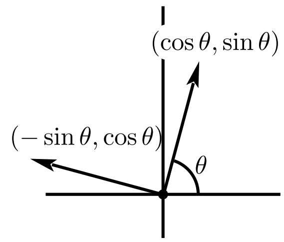


Fig 19

আবার দ্বিতীয় row-এর norm-এর square হল

$$\left\langle \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right), \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right) \right\rangle = \tfrac{1}{\sqrt{2}} \times \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}} \times \tfrac{1}{\sqrt{2}} = 1. \qquad (**)$$

লক্ষ কর যে (*) আর (**) আসলে নতুন কিছু নয়। $AA' = I$ দেখানোর সময়ে এই দুটো জিনিসই আমরা করেছিলাম। Column-এর ব্যাপারটাও একই রকম, ওটা তোমার নিজে নিজে করার জন্য রেখে দিলাম। ■

Exercise 17: এই দুটো vector নাও--

$$(\cos\theta, \sin\theta) \text{ এবং } (-\sin\theta, \cos\theta).$$

এদের ছবি রয়েছে Fig 19-তে। দেখাও যে, এদের norm হল 1, এবং এরা পরস্পরের সঙ্গে orthogonal. এবার এদেরকে পাশাপাশি লাগিয়ে এই matrix-টা বানাও $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. একে বলে **Given's rotation matrix.** এটা কি একটা orthogonal matrix? ■

এবার তবে প্রমাণ করা যাক। প্রথমে গুছিয়ে লিখে নিই কী প্রমাণ করতে চলেছি--

THEOREM

Let $A_{n\times n}$ be a real matrix. Then

it is orthogonal (unitary)

if and only if

its columns form an ONB of $\mathbb{R}^n$ ($\mathbb{C}^n$) under standard inner product,

if and only if

its rows form an ONB of $\mathbb{R}^n$ ($\mathbb{C}^n$) under standard inner product.

Proof: এখানে standard inner product মানে $\mathbb{R}^n$-এর বেলায় Euclidean inner product আর $\mathbb{C}^n$-এর বেলায় তারcomplex সংস্করণ। প্রমাণটা খালি orthogonal-এর জন্য করব। Unitary-র জন্য প্রমাণটা একেবারেই একইরকম।

Let $A_{n\times n} = [\ \vec{a}_1\ \cdots\ \vec{a}_n\ ]$. Then A is orthogonal iff

$$A'A = I,$$

iff

$$\begin{bmatrix} \vec{a}_1' \\ \vdots \\ \vec{a}_n' \end{bmatrix} [\ \vec{a}_1\ \cdots\ \vec{a}_n\ ] = I.$$

প্রথম অধ্যায়ে block matrix-এর গুণ শিখেছিলাম, মনে আছে? এবার সেটা কাজে লাগবে। এখানে row আর column-গুলো যেন একেকটা block. গুণফলের (i,j)-th entry-টা হবে $\vec{a}_i'\vec{a}_j$. এ থেকেই আসছে--

iff

$$\forall i,j \quad \vec{a}_i'\vec{a}_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$$

iff $\forall i \neq j \quad \langle \vec{a}_i, \vec{a}_j\rangle = 0$ and $\forall i,\ \|\vec{a}_i\| = 1$, as required.

Column-দের বেলায় প্রমাণটা হল। এবার row-দের পালা। অবশ্য নতুন করে কিছু করার নেই, transpose নিয়ে নিলেই চলবে--

If A is orthogonal, then so is A'. So considering A' in place of A, we get similar result about rows of A.

[Q.E.D]

Orthogonal matrix-দের একটা মস্ত গুণ হল ওদের দিয়ে গুণ করলে norm অপরিবর্তিত থাকে (standard inner product-এর ক্ষেত্রে)। এই কথাটাই অংকের ভাষায় লেখা হয়েছে নীচের theorem-এ।

THEOREM

Consider $\mathbb{R}^n$ with standard inner product. Let $A_{n\times n}$ be any orthogonal matrix. Then

$$\forall \vec{x} \in \mathbb{R}^n \quad \|A\vec{x}\| = \|\vec{x}\|.$$

Proof: প্রমাণ অত্যন্ত সহজ--

$$\begin{aligned} \|A\vec{x}\|^2 &= \langle A\vec{x}, A\vec{x}\rangle \\ &= (A\vec{x})'(A\vec{x}) \\ &= \vec{x}'A'A\vec{x} \\ &= \vec{x}'\vec{x} \quad \left[\because A \text{ is orthogonal}\right] \\ &= \|\vec{x}\|^2. \end{aligned}$$

> So the result is proved since norm ≥ 0.

[Q.E.D]

Exercise 18: একইরকম জিনিস খাটে unitary matrix-দের বেলাতেও। আমরা কাজ করব $\mathbb{C}^n$-এর উপরে standard inner product নিয়ে, মানে $\langle \vec{u}, \vec{v} \rangle = \vec{v}^* \vec{u}$. যদি $A_{n \times n}$ একটা unitary matrix হয়, তবে দেখাও যে--

$$\forall \vec{v} \in \mathbb{C}^n \quad \|A\vec{v}\| = \|\vec{v}\|.$$

■

আমরা দেখলাম যে, orthogonal/unitary matrix দিয়ে গুণ করলে vector-দের দৈর্ঘ্য অপরিবর্তিত থাকে। এরকম অন্য কোনো matrix কি আছে, যাদেরও একই ধর্ম? মানে যাদের দিয়ে গুণ করলেও দৈর্ঘ্য বদলায় না? উত্তর হল--না, নেই। অর্থাৎ যদি দেখো যে, $A_{n \times n}$ এমন একটা real (complex) matrix, যাকে দিয়ে গুণ করলে কখনোই দৈর্ঘ্য বদলায় না, অমনি জানবে যে, A অবশ্যই orthogonal (unitary) হতে বাধ্য! এর প্রমাণটা চেয়েছি নীচের অংকে।

Exercise 19: Let $A_{n \times n}$ be a real matrix such that $\forall \vec{x} \in \mathbb{R}^n \quad \|A\vec{x}\| = \|\vec{x}\|$. Show that A must be an orthogonal matrix.
HINT:
প্রথমে $\vec{x}$-এর জায়গায় $\vec{e}_i$ বসাও। তারপর $\vec{x}$-এর জায়গায় বসাও $\vec{e}_i + \vec{e}_j$. ■

THEOREM

An orthogonal/unitary matrix has all eigenvalues of modulus 1.

অর্থাৎ eigenvalue-গুলো 1, -1, i, $-i$, $\frac{1}{\sqrt{2}}(1+i)$, ইত্যাদি হতে পারে, কিন্তু 2 বা -3 বা $3-4i$ এইসব হতে পারে না।

Proof:

> Let A be an orthogonal/unitary matrix with an eigenvalue $\lambda \in \mathbb{C}$ and corresponding eigenvector $\vec{v}$.
>
> Then $A\vec{v} = \lambda\vec{v}$.
>
> Taking norms, $\|A\vec{v}\| = |\lambda| \cdot \|\vec{v}\|$.
>
> Since $\|A\vec{v}\| = \|\vec{v}\|$ and $\|\vec{v}\| \neq 0$, we have $|\lambda| = 1$, as required.

[Q.E.D]

যদি একটা orthogonal (unitary) matrix-এর eigenvalue-গুলোকে আমরা complex plane-এ বিন্দু হিসেবে এঁকে দেখাই, তবে বিন্দুগুলো সব থাকবে unit circle-এর পরিধির উপরে (Fig 20)। যদি matrix-টা real orthogonal হয়, তবে ব্যাপারটা Fig 21-এর মত আরেকটু সাজানো, real axis-এর উপরে-নীচে symmetric-ও বটে। এর কারণ eigenvalue-রা হল characteristic polynomial-এর zero-গুলো। যেহেতু characteristic polynomial-টা এখানে একটা real polynomial, তাই এই zero-গুলোর কেউ non-real হলেই, তার complex conjugate-টাও একটা zero হতে বাধ্য। এই symmetry-টা নিয়েই নীচের অংকটা।

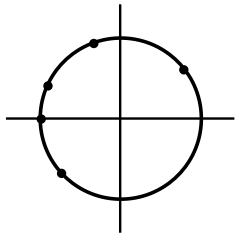


Fig 20

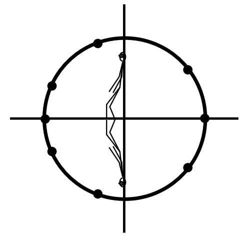


Fig 21

Example 14: Let P be a real orthogonal matrix with det $P = -1$. Prove that -1 is an eigenvalue of P.[3] **(2022.CC6.3gi)**

SOLUTION:

> Let $\lambda_1, ..., \lambda_n$ be all the eigenvalues (not necessarily distinct) of $P_{n\times n}$.
>
> Then we know that $\det(P) = \lambda_1 \times \cdots \times \lambda_n$.
>
> Since P is a real orthogonal matrix, $\forall i \quad |\lambda_i| = 1$. Also the non-real eigenvalues occur in conjugate pairs.

এখানেই symmetry-র কথাটা বললাম।

> So the product of all the non-real eigenvalues is 1.

এই লাইনটাই মোক্ষম ছিল। ধরো একটা eigenvalue হল $\lambda \notin \mathbb{R}$. তবে $\bar{\lambda}$-ও আরেকটা eigenvalue হবে। যদি এদের গুণ করি তবে পাব $\lambda\bar{\lambda} = |\lambda|^2 = 1$, যেহেতু $|\lambda| = 1$. এইভাবে প্রতি জোড়া non-real eigenvalue-দের গুণ করলেই 1 পাবে। এবার দেখি real eigenvalue-রা কী হতে পারে।

> The real eigenvalues are either 1 or -1.

কারণ real axis-টা unit circle-কে খালি দুটো বিন্দুতেই ছেদ করে, 1-এ আর -1-এ। তবেই বোঝো সবাইকেই গুণ করলে 1 হয়ে যাচ্ছে, একমাত্র ব্যতিক্রম হল -1. সুতরাং অন্ততঃ একটা -1 না থাকলে determinant-টা -1 হচ্ছে কী করে?

> Let, if possible, -1 be not an eigenvalue. Then the product of all the eigen values is $1 (\Rightarrow\Leftarrow)$.
>
> Hence, -1 must be an eigenvalue, as required.

■

Exercise 20: What can you say about the algebraic multiplicity of -1 in the above problem?
■

DAY 51
Gram-Schmidt process

আমরা আগেই বলেছি একটা basis নিয়ে কাজ করার সুবিধা কী। এর ফলে যেকোনো vector-কে একগুচ্ছ সংখ্যা দিয়ে প্রকাশ করা যায়। যেমন Fig 22-এ আমরা $\vec{v}$-কে $\{\vec{u}_1, \vec{u}_2\}$ basis দিয়ে এইভাবে লিখতে পারি--

$$\vec{v} = 2\vec{u}_1 + 3\vec{u}_2.$$

মানে $\{\vec{u}_1, \vec{u}_2\}$-এর সাপেক্ষে $\vec{v}$-কে লেখা যায় $(2,3)$ হিসেবে। এই কাজটা সব সময়েই করা সম্ভব বটে, কিন্তু যা খুশি একটা basis-এর সাপেক্ষে যেকোনো একটা vector-কে এইভাবে প্রকাশ করা সহজ নয়। কিন্তু যদি basis-টা একটা orthonormal basis (ONB) হয়, তবে কাজটা যে অত্যন্ত সহজ তা আমরা আগেই বলেছি। কারণ $\{\vec{u}_1, ..., \vec{u}_n\}$ একটা ONB হলে যেকোনো $\vec{v}$-কেই এভাবে লিখে ফেলা যায়--

$$\vec{v} = \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v}, \vec{u}_n \rangle \vec{u}_n.$$

সেই কারণে যেকোনো basis থেকে প্রথমে একটা ONB বার করে রাখলে সুবিধা হয়। এই কাজটা করার একটা কায়দা আছে, তাকে বলে **Gram-Schmidt process**, যেটা আমরা এবার শিখব।

51.1 জিনিসটা কী?

প্রথমে একটা উপমা মাথায় রাখলে সুবিধা হবে। Fig 23 দ্যাখো। এখানে তিনটে set দেওয়া আছে, যারা একজন আরেকজনের ঘাড়ে উঠে পড়েছে, মানে এদের মধ্যে intersection আছে। আমরা এদের বদলে এমন তিনটে set বানাতে চাই, যাদের মধ্যে কোনো intersection নেই, কিন্তু যাদের union এই তিনজনের union-এর সমান। অর্থাৎ এমন B_1, B_2, B_3 চাই, যাতে

$$B_1 \cup B_2 \cup B_3 = A_1 \cup A_2 \cup A_3$$

হয়, অথচ B_1, B_2, B_3-এর কারো সঙ্গে কারুর কোনো intersection না থাকে। সেই কাজটা ধাপে ধাপে এইভাবে করা যায়--

1. প্রথমে $B_1 = A_1$ নিই (Fig 24)।

2. এরপর তাকাই A_2-র দিকে। এর খানিকটা B_1-এর মধ্যে ঢুকে আছে। সেটুকু বাদ দিয়ে দিলে বাকি যেটুকু থাকে, সেটাকে বলি B_2, মানে B_2 হল Fig 25-এ চাঁদের মত অংশটা--

$$B_2 = A_2 \setminus (A_2 \cap B_1).$$

   এখানে '\' হল সেটের বিয়োগচিহ্ন।

Fig 22

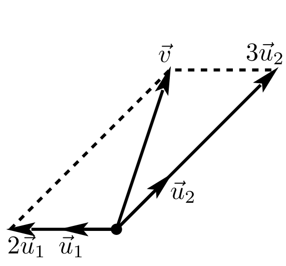


Fig 23

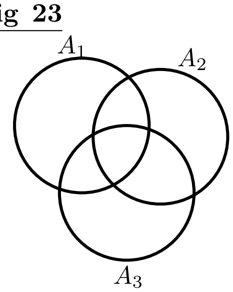


Fig 24

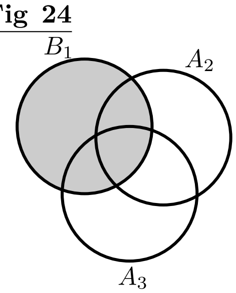
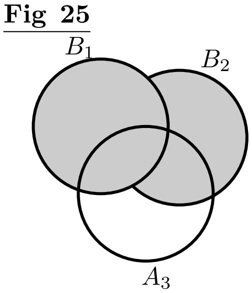


Fig 25



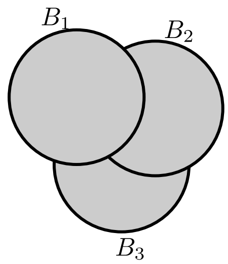


Fig 26

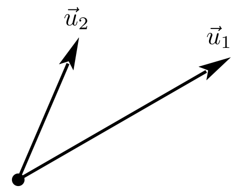


Fig 27

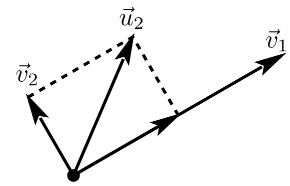


Fig 28

3. এবার A_3-র পালা। Fig 26 দেখো। এর যেটুকু B_1-এর মধ্যে সেটা বাদ দেব, যেটুকু B_2-র মধ্যে সেটাও বাদ যাবে। যেটুকু পড়ে থাকবে সেটাই নেব B_3, মানে

$$B_3 = A_3 \setminus \big[\underbrace{(A_3 \cap B_1)}_{\substack{\text{যেটুকু} \\ B_1\text{-এর} \\ \text{মধ্যে}}} \cup \underbrace{(A_3 \cap B_2)}_{\substack{\text{যেটুকু} \\ B_2\text{-এর} \\ \text{মধ্যে}}}\big].$$

ঠিক একই কাজ করা হয় Gram-Schmidt process-এও। প্রথমে খালি দুটো vector দিয়ে বোঝাই। Fig 27-এ $\mathbb{R}^2$-এর একটা basis রয়েছে, মানে দুটো vector, যারা একই লাইন বরাবর নয়। এটা থেকে একটা orthogonal basis বানাব, $\{\vec{v}_1, \vec{v}_2\}$. প্রথমে নেব

$$\vec{v}_1 = \vec{u}_1,$$

ঠিক যেমন $B_1 = A_1$ নিয়েছিলাম। তারপর দেখব $\vec{u}_2$-এর কতটা কাজ $\vec{v}_1$ করে দিতে পারে, মানে $\vec{v}_1$ বরাবর $\vec{u}_2$-এর component-টা কী। কী করে component বার করে সেটা গতকালই শিখেছি। সেই কায়দা লাগালে component-টা পাব--

$$\frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1.$$

এটা যেন আমাদের set-এর উপমায় $A_2 \cap B_1$-এর মত। এই অংশটা বাদ দিলে পড়ে থাকে যেটা, সেটাকেই নেব $\vec{v}_2$. বোঝার জন্য Fig 28 দেখো। সুতরাং--

$$\vec{v}_2 = \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1.$$

ঠিক যেমন set-এর উপমায় নিয়েছিলাম

$$B_2 = A_2 \setminus (A_2 \cap B_1).$$

এখানে যে $\langle \vec{v}_1, \vec{v}_1 \rangle$ দিয়ে ভাগ করলাম, তাতে কোনো অসুবিধা নেই, কারণ ওটা 0 হতে পারে না। $\vec{u}_1$ যেহেতু একটা basis-এর vector ছিল, তাই $\vec{u}_1 \neq \vec{0}$, ফলে $\vec{v}_1 = \vec{u}_1 \neq \vec{0}$, তাই $\langle \vec{v}_1, \vec{v}_1 \rangle \neq 0$.
সব মিলিয়ে আমাদের orthogonal basis-টা পেলাম $\{\vec{v}_1, \vec{v}_2\}$. এর থেকে ONB পাওয়া খালি norm দিয়ে ভাগ করার অপেক্ষা--

$$\left\{ \frac{\vec{v}_1}{\|\vec{v}_1\|}, \frac{\vec{v}_2}{\|\vec{v}_2\|} \right\}.$$

এখানে একটা ছোট্টো প্রশ্ন করে রাখি--

Exercise 21: বল তো কেন $\|\vec{v}_2\| \neq 0$ হবে? ■

যদি তিনটে vector নিয়ে শুরু করতাম $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, তবেও ঠিক একইভাবে এগোতাম--

$$\begin{aligned}
\vec{v}_1 &= \vec{u}_1 \\
\vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle}\vec{v}_1 \\
\vec{v}_3 &= \vec{u}_3 - \frac{\langle \vec{u}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle}\vec{v}_1 - \frac{\langle \vec{u}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle}\vec{v}_2.
\end{aligned}$$

এই $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ হবে একটা orthogonal basis. যদি orthonormal basis চাও, তবে $\vec{v}_i$-গুলোকে শেষে norm দিয়ে ভাগ করে দিলেই হবে--

$$\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}, \quad \vec{w}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|}, \quad \vec{w}_3 = \frac{\vec{v}_3}{\|\vec{v}_3\|}.$$

যদি n-খানা vector থাকত তবেও একইভাবে করা যেত--

Gram-Schmidt process

Let $\{\vec{u}_1, ..., \vec{u}_n\}$ be a linearly independent set of vectors. Let, for $k = 1, ..., n$,

$$\vec{v}_k = \vec{u}_k - \sum_{i=1}^{k-1} \frac{\langle \vec{u}_k, \vec{v}_i \rangle}{\langle \vec{v}_i, \vec{v}_i \rangle}\vec{v}_i$$

and

$$\vec{w}_k = \frac{\vec{v}_k}{\|\vec{v}_k\|}.$$

Then $\{\vec{v}_1, ..., \vec{v}_n\}$ is an orthogonal basis and $\{\vec{w}_1, ..., \vec{w}_n\}$ is an ONB for span$\{\vec{u}_1, ..., \vec{u}_n\}$.

Example 15: Using Gram-Schmidt orthonormalization process, find an orthonormal basis corresponding to the basis $\{(1,0,1), (1,0,-1), (0,3,4)\}$ in $\mathbb{R}^3(\mathbb{R})$ using standard inner product.[5] **(2022.CC12.3c)**

SOLUTION: প্রথমে সবকিছুর নাম দিয়ে নিই--

Let

$$\begin{aligned}
\vec{u}_1 &= (1,0,1), \\
\vec{u}_2 &= (1,0,-1), \\
\vec{u}_3 &= (0,3,4).
\end{aligned}$$

তারপর orthogonal বানাই--

Gram-Schmidt process:

$$\begin{aligned}
\vec{v}_1 &= \vec{u}_1 = (1,0,1)\\
\langle \vec{v}_1, \vec{v}_1 \rangle &= 2,\\
\langle \vec{u}_2, \vec{v}_1 \rangle &= 0,\\
\vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle}\vec{v}_1\\
&= (1,0,-1)\\
\langle \vec{v}_2, \vec{v}_2 \rangle &= 2,\\
\langle \vec{u}_3, \vec{v}_1 \rangle &= 4,\\
\langle \vec{u}_3, \vec{v}_2 \rangle &= -4,\\
\vec{v}_3 &= \vec{u}_3 - \frac{\langle \vec{u}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle}\vec{v}_1 - \frac{\langle \vec{u}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle}\vec{v}_2\\
&= (0,3,0).
\end{aligned}$$

এবার norm দিয়ে ভাগ করার পালা, যাকে বলে normalisation–

Normalisation:

$$\begin{aligned}
\vec{w}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} = \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)\\
\vec{w}_2 &= \frac{\vec{v}_2}{\|\vec{v}_2\|} = \left(\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}\right)\\
\vec{w}_3 &= \frac{\vec{v}_3}{\|\vec{v}_3\|} = (0,1,0).
\end{aligned}$$

Thus the required ONB for $\mathbb{R}^3$ is $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$.

■

একইরকম কিছু অংক তোমার জন্য।

Exercise 22: Apply Gram-Schmidt process to the vectors $(1,0,1)$, $(1,0,-1)$ and $(1,3,4)$ to obtain an orthonormal basis for $\mathbb{R}^3$ with the standard inner product. ■

Exercise 23: Use Gram-Schmidt process of orthogonalisation to construct an orthonormal basis for the subspace of $\mathbb{R}^4$ generated by $(1,1,0,1)$, $(1,-2,0,0)$, $(1,0,-1,2)$. ■

Exercise 24: Use Gram-Schmidt orthonormalization process to find an orthonormal basis of $\mathbb{R}^3$ from the basis $\{(1,0,1),(1,1,1),(1,3,4)\}$.[3] **(2020.CC12.3gi)** ■

Exercise 25: Use Gram-Schmidt process to obtain an orthonormal basis of the subspace of the Euclidean space $\mathbb{R}^3$ with standard inner product, generated by the linearly independent set

$$\{(1,1,1),\ (2,-2,1),\ (3,1,2)\}.$$

■

Exercise 26: Use Gram-Schmidt process to obtain an orthonormal basis of the subspace of the Euclidean space $\mathbb{R}^4$ with standard inner product, generated by the linearly independent set $\{(1,1,0,1),(1,1,0,0),(0,1,0,1)\}$. ■

51.2 Gram-Schmidt process-এর সুবিধা

যেকোনো একটা (finite) basis দিলেই সেটা থেকে Gram-Schmidt process দিয়ে একটা ONB বানিয়ে নেওয়া যায়। এতে অনেক সুবিধা। Basis-এর সম্পর্কে আমরা যা যা শিখেছিলাম, তার অনেক কিছুই Gram-Schmidt process-এর দৌলতে ONB-এর বেলাতেও খাটে। এবার সেরকম কিছু ব্যাপার আলোচনা করি।

51.2.1 ONB কি সব সময়েই থাকে?

আমরা জানি যে, যেকোনো finite dimensional vector space-এরই basis থাকে। স্বভাবতঃই প্রশ্ন ওঠে--যেকোনো finite dimensional inner product space-রই কি ONB-ও থাকে? Gram-Schmidt process-এর কথা ভাবলেই বুঝবে যে উত্তর হল--হ্যাঁ। কারণ যেই তুমি একটা basis পাচ্ছ, অমনি তার উপর Gram-Schmidt process লাগিয়ে দিলেই তো একটা ONB পেয়ে যাবে!

51.2.2 ONB-র জন্য extension theorem

আমরা শিখেছিলাম যে, একটা finite dimensional vector space-এর মধ্যে কোনো linearly independent set দেওয়া থাকলে সেটাকে একটা basis-এ extend করা যায় (ইঁদুর থেকে শুয়োর তৈরী করা)। এবার ধরো একটা orthogonal set দেওয়া আছে

$$\{\vec{v}_1,...,\vec{v}_k\},$$

যেখানে প্রত্যেকটা $\vec{v}_i\neq\vec{0}$. তাহলে কি এটাকে extend করে একটা orthogonal basis পাওয়া যাবে? উত্তর হল--হ্যাঁ। কারণটা এইরকম--

যেহেতু $\vec{v}_i$-রা পরস্পরের সাথে orthogonal, এবং সবাই nonzero, তাই $\{\vec{v}_1,...,\vec{v}_k\}$ একটা linearly independent set-ও বটে। সুতরাং extension theorem লাগিয়ে ওটা থেকে একটা basis পাওয়া যাবে--

$$\{\vec{v}_1,...,\vec{v}_k,\vec{v}_{k+1},...,\vec{v}_n\}.$$

এবার আমরা এর উপরে Gram-Schmidt process লাগাব। এখানে প্রথম k-খানা vector তো orthogonal আছেই, তাই Gram-Schmidt process-এর ফলে ওদের কোনো পরিবর্তন হবে না। সুতরাং শেষ পর্যন্ত যে orthogonal basis-টা পাব, সেটা দেখতে হবে এইরকম--

$$\{\vec{v}_1,...,\vec{v}_k,\vec{u}_{k+1},...,\vec{u}_n\},$$

যেটা $\{\vec{v}_1,...,\vec{v}_k\}$-এর একটা extension.

51.3 Parseval's identity এবং Bessel's inequality

ধরো তোমাকে $\mathbb{R}^3$-এর মধ্যে এই vector-টা দিলাম, $(2,3,4)$. এর Euclidean norm হবে $\sqrt{2^2+3^2+4^2}=\sqrt{29}$. এই vector-টার মধ্যে এই যে তিনটে সংখ্যা $2,3$ আর 4 এদের মিলিয়ে Euclidean basis-এর সাপেক্ষে vector-টার coordinate representation পাওয়া যায়।

এবার একটা নতুন basis দিচ্ছি, $\{(1,0,0),(1,1,0),(1,1,1)\}$. এর সাপেক্ষে ওই একই vector-এর coordinate representation হবে $(-1,-1,4)$, কারণ--

$$(2,3,4) = -(1,0,0) - (1,1,0) + 4(1,1,1).$$

এই coefficient-গুলোর square-দের যোগ করে square root নিলেও কি norm-টা পাওয়া যাবে? উত্তর হল--না, কারণ $\sqrt{(-1)^2 + (-1)^2 + 4^2} = \sqrt{18} \neq \sqrt{29}$.
তার মানে যেকোনো basis-এর সাপেক্ষে coordinate-এর square-গুলোর যোগফলের square root করে দিলেই norm নাও আসতে পারে। বস্তুতঃ এই গ্যারান্টিটা দেওয়া যায় কেবলমাত্র ONB-দের বেলাতেই। সেটাই আমাদের এবারের আলোচ্য বিষয়।
ধরো $\{\vec{v}_1, ..., \vec{v}_n\}$ হল একটা ONB. কোনো একটা vector দিলাম $\vec{v}$. আমরা যদি এটাকে ওই ONB দিয়ে লিখি, তবে এরকম কিছু একটা পাব--

$$\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n.$$

সেক্ষেত্রে $\sum |a_i|^2 = \|\vec{v}\|^2$ হবেই। এখানে $|a_i|^2$ লিখেছি, কারণ $\mathbb{C}$-এর বেলায় $|a_i|^2$ আর a_i^2 আলাদা। এটাকে বলে **Parseval's identity**. এটা কেন হবে, সেটা এক্ষুণি গুছিয়ে লিখে প্রমাণ করছি। খালি চট্ করে মনে করিয়ে দিই যে, $a_i = \langle \vec{v}, \vec{v}_i \rangle$.

Parseval's identity

Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be an ONB of an inner product space V. Then, for all $\vec{v} \in V$,

$$\|\vec{v}\|^2 = \sum_1^n |\langle \vec{v}, \vec{v}_i \rangle|^2.$$

Proof: প্রমাণটা অতি সংক্ষিপ্ত--

$$\begin{aligned}
\|\vec{v}\|^2 &= \langle \vec{v}, \vec{v} \rangle \\
&= \langle a_1\vec{v}_1 + \cdots + a_n\vec{v}_n, a_1\vec{v}_1 + \cdots + a_n\vec{v}_n \rangle \\
&= \sum_i a_i \bar{a}_i \langle \vec{v}_i, \vec{v}_i \rangle \quad [\because \forall i \neq j \quad \langle \vec{v}_i, \vec{v}_j \rangle = 0.] \\
&= \sum_i |a_i|^2 \quad [\because \forall i \quad \langle \vec{v}_i, \vec{v}_i \rangle = 1.]
\end{aligned}$$

[Q.E.D]

এর উল্টো দিকটাও সত্যি, যদিও সেটা প্রমাণ করাটা তত সহজ নয়--

Exercise 27: Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be a basis of a real inner product space V such that for all $\vec{v}$, if $\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$, then $\|\vec{v}\|^2 = \sum |a_i|^2$. Then prove that $\{\vec{v}_1, ..., \vec{v}_n\}$ must be an ONB.
HINT: প্রথমে দেখাও যে, প্রতিটা $\|\vec{v}_i\| = 1$. তারপর দেখাও $k \neq \ell$ হলে $\|\vec{v}_k + \vec{v}_\ell\| = \sqrt{2}$ হবে। সেখান থেকে দেখাও $\langle \vec{v}_k, \vec{v}_\ell \rangle = 0$. ■

এবার ধরো $\{\vec{v}_1, ..., \vec{v}_n\}$ হল একটা ONB, আর $\vec{v} \in V$. এখানেও যথারীতি

$$\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$$

লিখলাম। কিন্তু এবার $\|\vec{v}\|^2$-এর সঙ্গে তুলনা করব $\sum_1^k |a_i|^2$, যেখানে $k \leq n$. তবে কোন্টা বড় হবে? নিশ্চয়ই কোনো সন্দেহ নেই যে, $\|\vec{v}\|^2$-ই জিতবে, কারণ সেটা তো $\sum_1^n |a_i|^2$-এর সমান, যেটা n-খানা nonnegative সংখ্যার যোগফল, আর আমাদের sum-টাতে তো তার মধ্যে থেকে মোটে k-খানা সংখ্যা আছে! এটাই হল **Bessel's inequality**.

Bessel's inequality

Let $\{\vec{v}_1, ..., \vec{v}_k\}$ be any orthonormal set in an inner product space V. Then, for all $\vec{v} \in V$,

$$\sum_1^k |\langle \vec{v}, \vec{v}_i \rangle|^2 \leq \|\vec{v}\|^2.$$

Parseval's identity-র শর্তের সঙ্গে এখানকার শর্তটার পার্থক্য লক্ষ্য করো, ওখানে শুরু করেছিলাম একটা ONB নিয়ে, এখানে শুরু করেছি যেকোনো orthonormal set নিয়ে।

Proof: প্রমাণটা হবে Parseval's identity ব্যবহার করেই। আমরা $\vec{v}$-কে দুটো component-এ ভেঙে নেব, একটা থাকবে $\text{span}\{\vec{v}_1, ..., \vec{v}_k\}$ বরাবর, আর অন্যটা তার সঙ্গে orthogonal. কায়দাটা একেবারেই Gram-Schmidt process-এর মতই--

> Let $\vec{s} = \sum_1^k \langle \vec{v}, \vec{v}_i \rangle \vec{v}_i$ and $\vec{t} = \vec{v} - \vec{s}$.

আশা করি Gram-Schmidt process-এর সঙ্গে সাদৃশ্যটা চোখে পড়ছে। যেহেতু $\|\vec{v}_i\| = 1$, তাই $\langle \vec{v}_i, \vec{v}_i \rangle$ দিয়ে আর কষ্ট করে ভাগ করতে হয়নি।

> Then $\langle \vec{s}, \vec{t} \rangle = 0$.
> So $\|\vec{v}\|^2 = \|\vec{s} + \vec{t}\|^2 = \|\vec{s}\|^2 + \|\vec{t}\|^2$.

এবার $\vec{s}$-এর উপর Parseval's identity লাগাব--

> Now, by Parseval's identity, $\|\vec{s}\|^2 = \sum_1^k |\langle \vec{v}, \vec{v}_i \rangle|^2$.
> Thus,
> $$\|\vec{v}\|^2 = \|\vec{s}\|^2 + \|\vec{t}\|^2 \geq \sum_1^k |\langle \vec{v}, \vec{v}_i \rangle|^2,$$
> as required, since $\|\vec{t}\|^2 \geq 0$.

[Q.E.D]

Exercise 28: If $\{\beta_1, \beta_2, ..., \beta_r\}$ is an orthogonal set of vectors in an inner product space $V(\mathbb{R})$, then prove that for any vector $\alpha \in V$, $\|\alpha\|^2 \geq c_1^2 + c_2^2 + \cdots + c_r^2$, where c_i is the scalar component of α along β_i, $i = 1, 2, ..., r$.[3] **(2022.CC12.3aii)**

HINT: Bessel's inequality-র প্রমাণটাই চেয়েছে। ■

DAY 52
Projection

এবার আমরা projection বলে একটা জিনিস শিখব, যেটা linear algebra-র একটা অত্যন্ত গুরুত্বপূর্ণ ধারণা। অংকের বিভিন্ন শাখা তো বটেই, এমন কি বিজ্ঞানের অন্যান্য ক্ষেত্রেও এর প্রয়োগ আছে।

52.1 জিনিসটা কী?

> গাছের চেয়ে গাছের ছায়া সবরকমেই ভালো।
>
> --সুকুমার রায়

একটা উদাহরণ নিয়ে শুরু করি। ধরো একটা লোক রোদের মধ্যে দাঁড়িয়ে আছে (Fig 29)। মাটিতে তার ছায়া পড়েছে। আমরা বলব যে, ছায়াটা হল মাটিতে লোকটার projection. এই শব্দটা মোটেই অপরিচিত লাগার কথা নয়। সিনেমায় আমরা projector ব্যবহারের কথা জানি। সেখানে ফিল্মের ছবিগুলোকে project করা হয় পর্দায় (Fig 30)। সূর্যের উদাহরণটায় লোকটা যেন ফিল্মের উপরের ছবি, আর পর্দার ভূমিকা পালন করছে মাটিটা।

52.2 বিভিন্ন ধরণের projection

Projection নানা ধরণের হয়। সেটা নির্ভর করে দুটো জিনিসের উপর-- আলোটা কীভাবে পড়ছে, এবং পর্দাটা কীভাবে রয়েছে। যদি আলোটা একটা বিন্দু থেকে বেরোয় (যেমন সিনেমার projection-এর ক্ষেত্রে), তবে তাকে বলে **perspective projection.** যদি আলোর কিরণগুলো কোনো একদিক থেকে parallel-ভাবে (সমান্তরালভাবে) আসে (সূর্যের কিরণের মত), তবে তাকে বলে **parallel projection.** দুই ক্ষেত্রেই linear algebra কাজে লাগে। Parallel projection ব্যাপারটা বোঝা অপেক্ষাকৃত সহজতর। আমরা সেটাই শিখব। যদি আলোটা পর্দার উপর perpendicularly (লম্বভাবে) পড়ে, তবে অংকটা আরো সহজ হয়ে যায়, এবং এই জিনিসটারই প্রয়োগ সবচেয়ে বেশী। Parallel projection-এর এই বিশেষ ক্ষেত্রটাকে বলে **orthogonal projection.** এখানটায় একটু গোল বেঁধে যেতে পারে। Parallel (সমান্তরাল) আর orthogonal (লম্ব) পরস্পরবিরোধী বলে মনে হয়, যেমন দুটো লাইন যদি parallel হয় তবে তারা পরস্পরের সঙ্গে orthogonal হতে পারে না। কিন্তু projection-এর বেলায় orthogonal projection হল একধরণের parallel projection, কারণ--

- এখানে "parallel" মানে আলোর কিরণগুলো parallel,
- আর "orthogonal" মানে পর্দাটা আলোর কিরণের সঙ্গে লম্বভাবে রাখা।

একটা উদাহরণ নিই $\mathbb{R}^2$-তে। Fig 31-এর দিকে তাকাও। এখানে পর্দাটা রয়েছে x-axis বরাবর, আর আলোটা পড়ছে y-axis-এর সঙ্গে parallel-ভাবে। এবার একটা vector নিলাম যেমন ছবিতে দেখিয়েছি। তবে এর projection-টা কী হবে, মানে ছায়াটা কীভাবে পড়বে? উত্তরটা খুবই সহজ, Fig 32-এ দেখিয়েছি। লক্ষ্য কর, vector-টা এর ফলে দুভাগে ভেঙে গেল--

$$(x, y) = (x, 0) + (0, y).$$

এর মধ্যে প্রথম অংশটা রয়েছে পর্দার উপরে, এটাই হল projection-টা। বাকি অংশটা রয়েছে আলোর কিরণের সঙ্গে মিশে। আরেকটা উদাহরণ দেখি, এবার আলোটা ধরো বিকেলের পড়ন্ত রৌদ্র, একদিক থেকে হেলে পড়ছে (Fig 33)। এখানেও

Fig 29



Fig 30

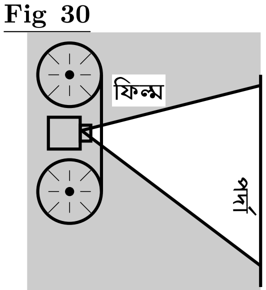
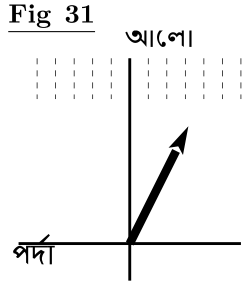


Fig 31



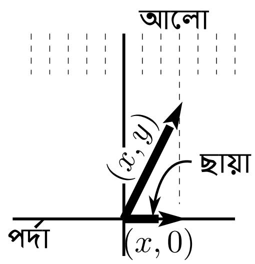


Fig 32

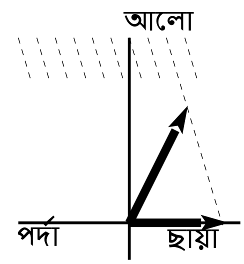


Fig 33

vector-টা দুভাগে ভেঙে যাচ্ছে, একটা ভাগ হল projection-টা স্বয়ং, আর অন্য ভাগটা মিশে আছে আলোর কিরণের সাথে। এই ব্যাপারটাকে আমরা এবার অংকের ভাষায় প্রকাশ করব।
প্রথম উদাহরণে পর্দাটা ছিল x-axis বরাবর। মানে, পর্দাটা হল $\mathbb{R}^2$-এর একটা subspace,

$$U = \{(x,0) \ : \ x \in \mathbb{R}\}.$$

আর আলোটা পড়ছে y-axis-এর সঙ্গে parallel-ভাবে। সুতরাং সেখান থেকে পাচ্ছি আরেকটা subspace

$$W = \{(0,y) \ : \ y \in \mathbb{R}\}.$$

লক্ষ্য কর যে, $\mathbb{R}^2$-র মধ্যে যাই (x,y) নিই না কেন, তাকে দুটো ভাগে ভেঙে ফেলা যাবে--

$$(x,y) = (x,0) + (0,y).$$

ঠিক যেন এই দুটো subspace মিলে পুরো $\mathbb{R}^2$-কে তৈরী করছে। সাবধান, এমন কিন্তু বলছি না যে, subspace-দুটোর union হল $\mathbb{R}^2$. খালি বলছি যে, ওদের দুজনকে যোগ করে পুরো $\mathbb{R}^2$-কে পাওয়া যায়, মানে

$$V = U + W.$$

আরও একটা জিনিস লক্ষ্য করো। U এবং V-এর intersection খালি একটাই বিন্দুতে, সেটা হল $(0,0)$. অর্থাৎ

$$U \cap W = \{\vec{0}\}.$$

আমরা দ্বিতীয় অধ্যায়েই দেখেছি যে, এই দুটো শর্তকে একসাথে লেখা হয়--

$$V = U \oplus W.$$

আর মুখে বলে--U আর W হল V-এর মধ্যে পরস্পরের direct complement.
ছায়া পড়ার গল্পটা যেকোনো direct complement জুটির বেলাতেই খাটে। জুটির মধ্যে একটাকে পর্দা বলে ভাবো, আর আলোর রশ্মিগুলোকে আসতে দাও অন্যটা বরাবর। অবশ্য dimension একটু বেশী হয়ে গেলে এভাবে কল্পনা করাটা দুষ্কর হয়ে পড়বে। তাই অংক দিয়ে প্রমাণ করে নিই। নীচের theorem-টা দ্বিতীয় অধ্যায়ে একটা অংক হিসেবে দিয়েছিলাম। পাছে সেটা না করে থাকো, তাই এখানে গুছিয়ে লিখে দিচ্ছি।

THEOREM

Let V be a vector space, and let $V = U \oplus W$. Then for every $\vec{v} \in V$ there is unique $\vec{u} \in U$ and unique $\vec{w} \in W$ such that $\vec{v} = \vec{u} + \vec{w}$.

Proof:

> Take any $\vec{v} \in V$.
> Since $V = U + W$, hence $\exists \vec{u} \in U, \vec{w} \in W \quad \vec{v} = \vec{u} + \vec{w}$.

Existence দেখানো হল, এবার uniqueness দেখানোর পালা--

> Let $\vec{u}_1 \in U$ and $\vec{w}_1 \in W$ be such that $\vec{v} = \vec{u}_1 + \vec{w}_1$.

তবে আমরা দেখাব যে, $\vec{u} = \vec{u}_1$ এবং $\vec{w} = \vec{w}_1$ হতে বাধ্য। Uniqueness দেখানো মানে যে এইটা দেখানো, সেটা ভালো করে বুঝে নাও।
আমাদের হাতে ছিল $\vec{v} = \vec{u} + \vec{w}$. এখন তার সাথে আরো ধরেছি $\vec{v} = \vec{u}_1 + \vec{w}_1$. সুতরাং--

> So $\vec{u} + \vec{w} = \vec{u}_1 + \vec{w}_1$, and hence $\vec{u} - \vec{u}_1 = \vec{w}_1 - \vec{w}$.
> Now $\vec{u} - \vec{u}_1 \in U$ and $\vec{w}_1 - \vec{w} \in W$.

যেহেতু ওরা পরস্পরের সমান, তাই ওরা দুজনেই U-তেও আছে, আবার W-তেও আছে, মানে $U \cap W$-তে আছে--

> So $\vec{u} - \vec{u}_1 = \vec{w}_1 - \vec{w} \in U \cap W = \{\vec{0}\}$.
> Thus, $\vec{u} = \vec{u}_1$ and $\vec{w} = \vec{w}_1$, as required.

[Q.E.D]

প্রমাণটা একেবারেই কঠিন কিছু নয়, তবু এখানে কেন U আর W-এর পক্ষে পরস্পরের direct complement হওয়া প্রয়োজন, সেটা একটু উদাহরণ দিয়ে বুঝে নেওয়া যাক। Direct complement-এর সংজ্ঞায় দুটো শর্ত আছে--এক, $V = U + W$, আর দুই, $U \cap W = \{\vec{0}\}$. যদি $V = U + W$ না হত, তবে ছায়াটার কী সমস্যা হত? একটা উদাহরণ নিই। ধরো $V = \mathbb{R}^3$ আর U হল x-axis. মানে এখানে পর্দাটা আসলে একটা সূতোর মত x-axis বরাবর খাটানো আছে। W নিলাম মনে করো z-axis. অর্থাৎ আলোটা পড়ছে খাড়াভাবে। এবার যদি একটা vector নিই $(1, 2, 3)$, তবে তার ছায়া পর্দায় কোথায় পড়বে? উত্তর হল--কোথাও না, vector-টার ডগা ছুঁয়ে যে কিরণটা নামছে, তার সঙ্গে x-axis-এর কোনো দিন দেখাই হবে না (Fig 34). এখানে পর্দাটা খাটানো উচিত ছিল পুরোটা জুড়ে, যাতে সব আলোকরশ্মিই পর্দায় পড়তে বাধ্য হয়। এবার বুঝলে $V = U + W$ কেন লাগছে? অন্য শর্তটা, মানে $U \cap W = \{\vec{0}\}$ কেন দরকার, সেটাও ছবি দিয়ে সহজে বোঝা যায়। মনে করো U নিলাম পুরো xy-plane-টা। আর W নিলাম x-axis. এবারও সেই একই সমস্যা, $(1, 2, 3)$-র কোনো ছায়া পড়বে না (Fig 35).
এবার তবে অংকের ভাষায় projection-এর সংজ্ঞা লিখে ফেলতে আর কোনো অসুবিধা নেই। সংজ্ঞাটা আসছে উপরের theorem-টা থেকে, তাই সংজ্ঞার গোড়াতেই theorem-টা আরেকবার বলে নিতে হয়েছে।

Fig 34

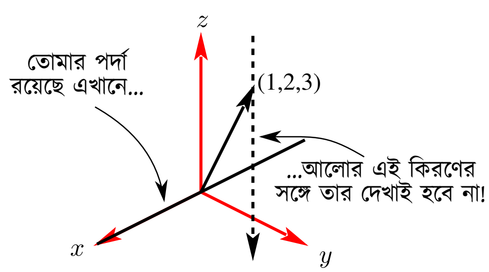
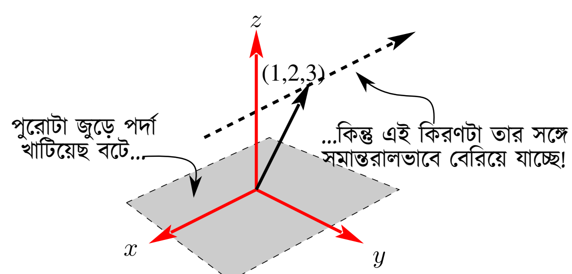


Fig 35



> DEFINITION: **Projection**
>
> Let $V = U \oplus W$. Then for any $\vec{v} \in V$ we have unique $\vec{u} \in U$ and $\vec{w} \in W$ such that $\vec{v} = \vec{u} + \vec{w}$. This $\vec{u}$ is called the **projection of $\vec{v}$ onto U along W.**

অবশ্যই তুমি একই সংজ্ঞা অনুযায়ী $\vec{w}$-কে বলতে পারো projection of $\vec{v}$ onto W along U. এখানে U আর W-এর মধ্যে কোনটাকে পর্দা বলে ভাববে আর কোনটাকে আলোর কিরণের অভিমুখ ভাববে, সেটা সম্পূর্ণই তোমার অভিরুচি।

52.3 Orthogonal complement

আগেই বলেছি যে, বিভিন্ন ধরণের parallel projection-এর মধ্যে সবচেয়ে বেশী ব্যবহার হয় orthogonal projection, মানে যেখানে আলোর কিরণগুলো লম্বভাবে এসে পর্দার উপরে পড়ে। তার মানে এখানে আমাদের এমন দুটো direct complement নিতে হবে U আর W, যাতে যাই $\vec{u} \in U$ আর $\vec{w} \in W$ নিই না কেন, সবসময়েই $\langle \vec{u}, \vec{w} \rangle = 0$ হবে। এরকম direct complement-দের বলে **orthogonal complement**.

> DEFINITION: **Orthogonal complement**
>
> Let V be any inner product space. Then two subspaces U and W are called **orthogonal complement**s of each other in V if
>
> $$V = U + W$$
>
> and
>
> $$\forall \vec{u} \in U \quad \forall \vec{w} \in W \quad \langle \vec{u}, \vec{w} \rangle = 0.$$
>
> Here we write $U = W^{\perp}$ and $W = U^{\perp}$.

আমরা পড়ার সময়ে $W^{\perp}$-কে পড়ি "W perp."
লক্ষ্য কর যে, এখানে আর $U \cap W = \{\vec{0}\}$ শর্তটা আলাদা করে লিখতে হল না। কারণ, যদি $\vec{v} \in U \cap W$ হয়, তবে $\vec{v} \in U$ এবং $\vec{v} \in W$ হওয়ায় $\langle \vec{v}, \vec{v} \rangle = 0$ হতে বাধ্য, মানে $\|\vec{v}\|^2 = 0$, ফলে $\vec{v} = \vec{0}$ হবে।
দেখতেই পাচ্ছ যে, দুটো subspace পরস্পরের orthogonal complement হলে তারা পরস্পরের direct complement-ও বটে। Direct complement-এর ধারণাটা বেশী general, সেটা যে কোনো vector space-এর বেলাতেই খাটে। কিন্তু orthogonal complement নিয়ে আলোচনা করার জন্য একটা inner product space নিয়ে কাজ করতে হয়। যদি n-dimensional কোনো inner product space-এর মধ্যে একটা k-dimensional subspace নাও, তবে তার dimension হবে $n-k$. এটা বস্তুতঃ যেকোনো direct complement-এর জন্যই খাটে।
Orthogonal complement একধরণের direct complement হলেও ওর একটা বাড়তি সুবিধা আছে। Direct complement-রা সাধারণতঃ unique হয় না, কিন্তু orthogonal complement সবসময়ে unique হয়। সেটা প্রথমে ছবি দিয়ে বুঝে নিই, তারপর অংক কষে প্রমাণ করব।

Example 16: ধরো $V = \mathbb{R}^2$ আর $U = \{(x, 0) \ : \ x \in \mathbb{R}\}$ নিলাম, মানে U হল x-axis-টা। এর দুটো direct complement দাও। $U^{\perp}$ কী হবে? এখানে আমরা standard inner product নিয়ে কাজ করছি।
SOLUTION: Direct complement-এর জন্য x-axis বাদে আর যেকোনো একটা লাইন নাও origin দিয়ে, যেমন ধরো Fig 36-এ নিয়েছি $(1,1)$ দিয়ে যে লাইনটা যায়, মানে $W = \{(x, x) \ : \ x \in \mathbb{R}\}$. তাহলে W হবে U-এর একটা direct complement. কেন হবে সেটা বোঝা সহজ। এরকম যেকোনো লাইনই একটা subspace, আর সেটা U-কে খালি origin-এ ছেদ করে। সুতরাং $U \cap W = \{\vec{0}\}$ হচ্ছে। আর U হল $\text{span}\{(1,0)\}$ এবং $W = \text{span}\{(1,1)\}$.যেহেতু $(1,0)$ আর $(1,1)$ একই লাইন বরাবর নয়, তাই ওরা পরস্পরের সঙ্গে linearly independent. এদিকে $\mathbb{R}^2$-এর dimension

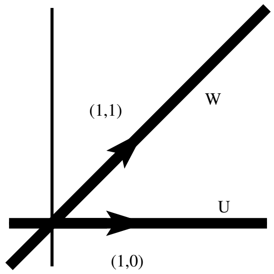


Fig 36

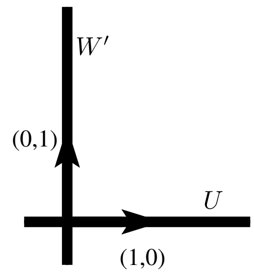


Fig 37

তো 2. সুতরাং $\{(1,0),(1,1)\}$ একটা basis. সুতরাং $\mathbb{R}^2 = U + W$-ও হচ্ছে। তাই W হবে U-এর একটা direct complement.
যদি অন্য কোনো লাইন নিই, যেমন ধরো y-axis-টা, মানে $W' = \text{span}\{(0,1)\}$, তাহলে সেটাও একই যুক্তিতে U-এর আরেকটা direct complement হবে। Fig 37 দেখো।
কিন্তু যদি orthogonal complement চাই, তবে বুঝতেই পারছ যে, y-axis-টা নেওয়া ছাড়া পথ নেই, কারণ y-axis-টাই origin দিয়ে যাওয়া একমাত্র লাইন, যেটা x-axis-এর সঙ্গে orthogonal-ভাবে আছে। ■

Example 17: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). The orthogonal complement of $W = \{(x,y,z) \in \mathbb{R}^3 \ : \ x+y-z = 0 \text{ and } x-2y+z = 0\}$ in $\mathbb{R}^3$ is

(A) $\{(x,y,z) \in \mathbb{R}^3 \ : \ x+y+z = 0\}$

(B) $\{(x,y,z) \in \mathbb{R}^3 \ : \ x+2y+3z = 0\}$

(C) $\{(x,y,z) \in \mathbb{R}^3 \ : \ \frac{x}{1} = \frac{y}{2} = \frac{z}{3}\}$

(D) None of the above.

(2020.CC12.1f)
SOLUTION: এরকম ক্ষেত্রে প্রথমেই dimension-টা মনে মনে মিলিয়ে নেওয়া ভালো। এখানে কাজ হচ্ছে $\mathbb{R}^3$-এর মধ্যে, যার dimension হল 3. এর উপর দুটো শর্ত চাপিয়ে W পেয়েছি, তাই[3] $\dim(W) = 3-2 = 1$. সুতরাং $\dim(W^{\perp}) = 3 - \dim(W) = 2$. দেখাই যাচ্ছে (C)-তে dimension হল $3-2=1$. সুতরাং ওটা হবে না। অবশ্য মোটে একটা option দূর করে খুব একটা লাভ হল না। কিন্তু তাও এই dimension-এর যুক্তিটা শিখে রাখা ভালো, অভ্যাস হয়ে গেলে ওটা চোখে দেখেই মুহূর্তের মধ্যেই লাগানো যায়। এবার খুঁটিনাটির দিকে তাকাই। একটা ভালো কায়দা হল W-এর সদস্যদের সহজ করে লিখে নেওয়া। এখানে $z = x+y$, আর তাই $x - 2y + (x+y) = 0$, অর্থাৎ $y = 2x$. সুতরাং $z = x+y = 3x$. তাই W-এর যেকোনো সদস্যই হল $(x,y,z) = (x,2x,3x)$ আকারের, এবং এই আকারের সব vector-ই W-এর মধ্যে আছে। তার মানে $W = \text{span}\{(1,2,3)\}$. সুতরাং $W^{\perp}$-এর সদস্যরা এমন (x,y,z) হবে, যাতে $\langle (x,y,z),(1,2,3)\rangle = 0$ হয়, মানে $x+2y+3z = 0$ হয়। তাই উত্তর হবে (B). ■

Exercise 29: Find the orthogonal complement of

$$W = \{(x,y,z) \in \mathbb{R}^3 \ : \ x+y = 0 \text{ and } x+z = 0\}$$

[3]শর্তদুটো linearly independent.

in $\mathbb{R}^3$. ■

Example 18: Let P_3 be the inner product space of all real polynomials of degree ≤ 3 with the inner product $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$; $f, g \in P_3$, and also let W be the subspace of P_3 with basis $\{1, t^2\}$. Find a basis for $W^{\perp}$.[5] **(2022.CC12.3b)**

SOLUTION:

> Here $W^{\perp}$ is the set of all polynomials of the form $p(t) = a + bt + ct^2$ such that
> $$\langle 1, p(t)\rangle = \left\langle t^2, p(t)\right\rangle = 0,$$
> ie,
> $$\int_0^1 1 \cdot p(t)\,dt = \int_0^1 t^2 \cdot p(t)\,dt = 0.$$
> So
> $$a + \frac{b}{2} + \frac{c}{3} = \frac{a}{3} + \frac{b}{4} + \frac{c}{5} = 0.$$

এই সম্পর্কগুলো ব্যবহার করে আমরা a, b, c-কে কোনো একটা vector দিয়ে লিখব। যেমন ধরো, c দিয়ে--

> So $a = -\frac{b}{2} - \frac{c}{3}$, and hence $-\frac{b}{6} - \frac{c}{9} + \frac{b}{4} + \frac{c}{5} = \frac{b}{12} + \frac{4c}{45} = 0$.
> Hence $b = -\frac{16c}{15}$, and so $a = \cdots = \frac{c}{5}$.
> So the general form of $p(t)$ is $p(t) = \frac{c}{15}(3 - 16t + 15t^2)$, where $c \in \mathbb{R}$ is arbitrary.
> Hence $\{3 - 16t + 15t^2\}$ spans $W^{\perp}$. Since $3 - 16t + 15t^2$ is a non-zero polynomial, hence $\{3 - 16t + 15t^2\}$ is also linearly independent, and so a basis of $W^{\perp}$, as required.

■

একইরকম একটা অংক তোমার জন্য।

Exercise 30: Let P_3 be the inner product space of all real polynomials of degree ≤ 3 with the inner product $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$; $f, g \in P_3$, and also let W be the subspace of P_3 with basis $\{1, t\}$. Find a basis for $W^{\perp}$. ■

Example 19: Choose the correct alternative with proper justification (1 mark for correct answer and 1 mark for justification). Let $\{u_1, u_2, u_3, u_4, u_5\}$ be an orthogonal basis of $\mathbb{R}^5$, $\vec{y}$ be a vector in $\mathbb{R}^5$, $W_1 = \text{span}\{u_1, u_2\}$, $W_2 = \text{span}\{u_3, u4, u_5\}$. If $W^{\perp}$ denotes the orthogonal complement of W, then which of the following is false?

(A) $W_1 = W_2^{\perp}$

(B) $W_2 = W_1^\perp$

(C) there are two vectors Z_1 in W_1 and Z_2 in W_2 such that $y = Z_1 + Z_2$.

(D) y is orthogonal to W_1 and to W_2.

(2021.CC12.1g)
SOLUTION:এখানে প্রথম তিনটে option সঠিক। শেষের option-টা ঠিক না ভুল সেটা y-টা কী, তার উপর নির্ভর করে। যেমন যদি $\vec{y} = \vec{0}$ হয়, তবে ঠিক, অন্য যেকোনো $\vec{y}$-এর বেলাতেই ভুল। তাই অগত্যা এখানে (D)-ই নির্বাচন করতে হবে। ■

এই প্রসঙ্গে আরেকটা দরকারী সংজ্ঞাও বলে নিই--

DEFINITION: **(Mutually) orthogonal subspaces**

Let V be an inner product space, and let U, W be two subspaces of V. Then we call U and W **mutually orthogonal subspaces** if

$$\forall \vec{u} \in U \quad \forall \vec{w} \in W \quad \langle \vec{u}, \vec{w} \rangle = 0.$$

এখানে "mutually" শব্দটা ব্র্যাকেটের মধ্যে রেখেছি, কারণ অনেক সময়েই ওটার উল্লেখ করা হয় না।
এবারের অংকটা একেবারেই সোজা।

Exercise 31: If U, W are mutually orthogonal subspaces then show that $U \cap W = \{\vec{0}\}$. ■

Exercise 32: Which of the following statements is/are true in $\mathbb{R}^3$ with standard inner product?

(A) The xz-plane and the xy-plane are mutually orthogonal subspaces.

(B) The x-axis and the y-axis are mutually orthogonal subspaces.

(C) Every subspace in $\mathbb{R}^3$ is an orthogonal complement of $\{\vec{0}\}$.

(D) If U, V and W are nontrivial, orthogonal subspaces of $\mathbb{R}^3$, then $(U \oplus V)^\perp = W$.

■

52.4 Orthogonal projection

ধরো একটা inner product space আছে V. এর মধ্যে কোনো একটা subspace দেওয়া আছে U, যার orthogonal complement হল W-ও জানা আছে। তবে যেকোনো $\vec{v} \in V$-কেই এইভাবে লেখা যাবে--

$$\vec{v} = \vec{u} + \vec{w},$$

যেখানে $\vec{u} \in U$ এবং $\vec{w} \in W$. এবং এইভাবে ভাঙাটা unique. এখানে $\vec{u}$-কে বলব $\vec{v}$-এর **orthogonal projection** onto U.

যদি U-টা finite dimensional হয়, তবে এই orthogonal projection-টা বার করার একটা সহজ কায়দা আছে, যেটা খুবই কাজে দেয়। ম্যাজিক দেখানোর জন্য যেমন আগেভাগে কিছু প্রস্তুতি করে রাখলে ভালো হয়, এখানেও তেমনি একটা প্রস্তুতি লাগে--subspace-টার একটা orthonormal basis বার করে রাখতে হয়। ধরো basis-টা হল--

$$\{\vec{u}_1, ..., \vec{u}_k\}.$$

এবার তোমাকে একটা কোনো $\vec{v} \in V$ দিয়ে বলা হল U-এর মধ্যে এর orthogonal projection বার কর। অমনি তুমি প্রত্যেকটা $\vec{u}_i$-এর সঙ্গে $\vec{v}$-এর inner product বার করে ফেলবে, মানে $\langle \vec{v}, \vec{u}_i \rangle$. এই ভাবে যে p-খানা সংখ্যা পাবে তাদের দিয়ে $\vec{u}_i$-গুলোর linear combination বানালেই উত্তর পেয়ে যাবে--

$$\langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v}, \vec{u}_p \rangle \vec{u}_p.$$

এটাই কেন U-এর মধ্যে $\vec{v}$-এর orthogonal projection হবে, সেটা তোমার ভেবে দেখার জন্য রেখে দিলাম।

52.5 Orthogonal complement-এর unique existence-এর প্রমাণ

একটা inner product space-এর মধ্যে যেকোনো finite dimensional subspace-এর orthogonal complement সর্বদা exist করে, এবং unique হয়। এবার সেটা প্রমাণ করা যাক।

> **THEOREM**
>
> Let V be any inner product space. Let U be any finite dimensional subspace of V. Then there is a unique subspace W of V such that W is orthogonal complement of U. Also, this W is given by
>
> $$W = \{\vec{v} \in V \ : \ \forall \vec{u} \in U \quad \langle \vec{v}, \vec{u} \rangle = 0\}.$$

এই theorem-টা এক ঢিলে তিন পাখি মেরেছে--

- এক, এ থেকে জানা যাচ্ছে যে, একটা inner product space-এর মধ্যে সব finite dimensional subspace-এরই orthogonal complement থাকে,
- দুই, সেটা unique হয়,
- আর তিন, সেটা কীরকম দেখতে, সেটাও পাওয়া যাচ্ছে। U-এর প্রত্যেক সদস্যর সঙ্গেই orthogonal যত vector আছে, তাদের set-টাই আসলে U-এর orthogonal complement!

এবার প্রমাণে হাত দিই।

Proof:

> Let W be as given.
>
> Step 1: Shall show that W is a subspace of V.
>
> Clearly $\vec{0} \in W$, since $\vec{0}$ is orthogonal to all vectors.

এবার দেখাতে হবে যে, যাই $\vec{w}_1, \vec{w}_2 \in W$ আর যাই দুটো scalar a, b নাও না কেন, $a\vec{w}_1 + b\vec{w}_2 \in W$ হবেই।

Take any $\vec{w}_1, \vec{w}_2 \in W$ and scalars a, b.
Then $\forall \vec{u} \in U \quad \langle \vec{w}_1, \vec{u} \rangle = \langle \vec{w}_2, \vec{u} \rangle = 0$.
So $\forall \vec{u} \in U \quad \langle a\vec{w}_1 + b\vec{w}_2, \vec{u} \rangle = a \langle \vec{w}_1, \vec{u} \rangle + b \langle \vec{w}_1, \vec{u} \rangle = 0$.
So $a\vec{w}_1 + b\vec{w}_2 \in W$, as required.
Step 2: By construction

$$\forall \vec{w} \in W \quad \forall \vec{u} \in U \quad \langle \vec{w}, \vec{u} \rangle = 0.$$

Step 3: Shall show $V = U + W$, ie,

$$\forall \vec{v} \in V \quad \exists \vec{u} \in U \quad \exists \vec{w} \in W \quad \vec{v} = \vec{u} + \vec{w}.$$

Take any $\vec{v} \in V$.
Let $\{\vec{u}_1, ..., \vec{u}_k\}$ be any ONB of U.
Choose $\vec{u} = \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v}, \vec{u}_k \rangle \vec{u}_k \in U$ and $\vec{w} = \vec{v} - \vec{u}$.

এই $\vec{u}$-টা আসলে U-এর মধ্যে $\vec{v}$-এর orthogonal projection, আর $\vec{w}$ হল যেটুকু projection করার পর পড়ে রইল। আমরা দেখাব $\vec{w} \in W$ হবে, মানে $\vec{w}$-টা U-এর সব vector-এর সঙ্গেই orthogonal হবে।

Then $\vec{w} \in W$.

⟦ Because:
Enough to show that $\langle \vec{w}, \vec{u}_i \rangle = 0$ for $i = 1, ..., k$.
Now

$$\begin{aligned} \langle \vec{w}, \vec{u}_i \rangle &= \left\langle \vec{v} - \sum_{j=1}^{k} \langle \vec{v}, \vec{u}_j \rangle \vec{u}_j, \vec{u}_i \right\rangle \\ &= \langle \vec{v}, \vec{u}_i \rangle - \langle \vec{v}, \vec{u}_i \rangle \langle \vec{u}_i, \vec{u}_i \rangle \quad [\because \langle \vec{u}_j, \vec{u}_i \rangle = 0 \text{ if } i \neq j] \\ &= 0 \quad [\because \langle \vec{u}_i, \vec{u}_i \rangle = 1]. \end{aligned}$$

⟧

Also $\vec{v} = \vec{u} + \vec{w}$, as required.

[Q.E.D]

এই প্রমাণটা ভালো করে পড়ে নিয়ে নীচের অংকটা কর।

Exercise 33: Let V be any inner product space. Let $S \subseteq V$ be any *subset* (not necessarily a *subspace*). Then show that

$$W = \{\vec{v} \in V \ : \ \forall \vec{u} \in S \quad \langle \vec{u}, \vec{v} \rangle = 0\}$$

is a subspace of V. ■

52.6 $\mathscr{R}(A)^\perp = \mathscr{N}(A)$

এই ধরো একটা matrix দিয়ে একটা vector-কে গুণ করলাম--

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 10 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

চট্ করে দেখে নাও ঠিক করেছি কিনা। এই যে উত্তরটা $(0, 0)$ হল, এর মধ্যে উপরের 0-টা পেয়েছি এইভাবে--

$$1 \times 10 + 2 \times (-5) = 0.$$

লক্ষ্য কর, এখানে A-র প্রথম row-টাই খালি কাজে লাগল। আমরা এটাকে এভাবেও লিখতে পারতাম--

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 10 \\ -5 \end{bmatrix} = 0 \text{ বা } \vec{b}_1'\vec{x} = 0,$$

যেখানে $\vec{b}_1$ হল matrix-টার প্রথম row. এবার লক্ষ্য করো যে, এটাকে আমরা Euclidean inner product দিয়ে ভাবতে পারি--

$$\vec{b}_1'\vec{x} = \left\langle \vec{b}_1, \vec{x} \right\rangle.$$

সুতরাং আমরা পেলাম--

$$\left\langle \vec{b}_1, \vec{x} \right\rangle = 0,$$

অর্থাৎ $\vec{b}_1$ এবং $\vec{x}$ দেখা যাচ্ছে পরস্পরের সঙ্গে orthogonal. একইভাবে দ্বিতীয় row-টার জন্য লিখতে পারি--

$$\left\langle \vec{b}_2, \vec{x} \right\rangle = 0.$$

অভ্যাস হয়ে গেলে পুরো ব্যাপারটা এক লাইনে লিখে দেওয়া যায়--

$$A\vec{x} = \begin{bmatrix} \vec{b}_1' \\ \vec{b}_2' \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{b}_1'\vec{x} \\ \vec{b}_2'\vec{x} \end{bmatrix} = \begin{bmatrix} \left\langle \vec{b}_1, \vec{x} \right\rangle \\ \left\langle \vec{b}_2, \vec{x} \right\rangle \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

এবার বল তো $\vec{x}$ কি $\mathscr{N}(A)$-এর মধ্যে আছে, নাকি বাইরে? অবশ্যই মধ্যে আছে, কারণ $A\vec{x} = \vec{0}$. তার মানে লক্ষ্য কর যে, এই দুটো জিনিস সমার্থক--

1. $\vec{x} \in \mathscr{N}(A)$.
2. $\vec{x}$-টা A-র প্রতিটা row-এর সঙ্গেই orthogonal.

এদিকে A-র প্রত্যেকটা row-এর সঙ্গে orthogonal হওয়া মানে $\mathscr{R}(A)$-র সবার সঙ্গেই orthogonal হওয়া। কারণ--

> যদি যেকোনো $\vec{b} \in \mathscr{R}(A)$ নাও, তবে $\vec{b}$-কে A-র row-দের linear combination হিসেবে লিখতে পারবে--
>
> $$\vec{b} = c_1\vec{b}_1 + \cdots + c_m\vec{b}_m,$$
>
> যেখানে c_i-গুলো হল কিছু scalar. আমরা জানি যে $\left\langle \vec{b}_i, \vec{x} \right\rangle = 0$. সুতরাং
>
> $$\left\langle \vec{b}, \vec{x} \right\rangle = c_1 \left\langle \vec{b}_1, \vec{x} \right\rangle + \cdots + c_m \left\langle \vec{b}_m \vec{x} \right\rangle = 0.$$

সুতরাং পেলাম

$$\mathscr{N}(A) = \mathscr{R}(A)^{\perp}.$$

অবশ্য চাইলে তুমি এটাকে ঘুরিয়ে এভাবেও লিখতে পারো--

$$\mathscr{R}(A) = \mathscr{N}(A)^{\perp}.$$

এটাকে একটা theorem-এর আকারে লিখে ফেলা যাক--

> **THEOREM**
>
> For any matrix A over the real field, the subspaces $\mathscr{R}(A)$ and $\mathscr{N}(A)$ are orthogonal complements of each other under the standard inner product.

মনে আছে নিশ্চয়ই যে, যেকোনো matrix-এর বেলাতেই rank এবং nullity যোগ করলে column-এর সংখ্যা পাওয়া যায়, অর্থাৎ যেকোনো $A_{m\times n}$-এর জন্যই--

$$\dim(\mathscr{R}(A)) + \dim(\mathscr{N}(A)) = n.$$

কিন্তু এই theorem-টা বলছে যে, real field-এর বেলায় ব্যাপারটা আরো গভীর। এক্ষেত্রে $\mathscr{R}(A)$ এবং $\mathscr{N}(A)$ দুজনে হল $\mathbb{R}^n$-এর মধ্যে পরস্পরের orthogonal complement.

52.6.1 একটা প্রয়োগ

Orthogonal projection-এর একটা প্রয়োগ দেখা যাক। ধরো একটা লাইন আছে Fig 38-এর মত। একটা point নিলাম P. যদি বলি লাইনের উপর কোন্ point-টা P-এর সবচেয়ে কাছে আছে, তবে তুমি P থেকে লাইনটার উপর perpendicular টানবে PQ, যেমন দেখিয়েছি Fig 39-এ। তাহলে Q-ই হবে লাইনটার উপরে P-এর নিকটতম বিন্দু। যদি প্রমাণ করতে হয়, অমনি তুমি লাইনের উপর যেকোনো বিন্দু নেবে R, এবং ΔPQR-এর দিকে তাকাবে। এটা একটা right angled triangle (সমকোণী ত্রিভুজ) সুতরাং $PQ \leq PR$ হতে বাধ্য। এখানে আসলে তুমি Pythagoras লাগালে--

$$PQ = \sqrt{PR^2 - QR^2} \leq \sqrt{PR^2} = PR.$$

ঠিক একই যুক্তি কাজ করে যেকোনো orthogonal projection-এর ক্ষেত্রেও--ধরো V হল একটা inner product space, এবং U হল তার কোনো একটা subspace. তোমাকে যা খুশি একটা $\vec{v} \in V$ দিলাম। তোমার কাজ হল এমন $\vec{u} \in U$ বার করা, যাতে $\|\vec{v} - \vec{u}\|$ হয় সবচেয়ে কম। তবে উত্তর হবে--

$\vec{u} =$ orthogonal projection of $\vec{v}$ onto U.

Fig 38

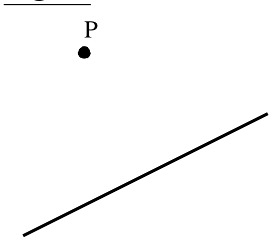
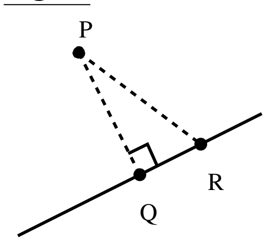


Fig 39



প্রমাণটা এখানেও একই কায়দায় হবে। যেকোনো একটা $\vec{u}_1 \in U$ নাও, এবং $\vec{v}, \vec{u}_1$ আর $\vec{u}$ দিয়ে তৈরী triangle-টার উপর Pythagoras খাটবে--

$$\|\vec{v} - \vec{u}_1\|^2 = \|\vec{v} - \vec{u}\|^2 + \|\vec{u} - \vec{u}_1\|^2 . \qquad (*)$$

তাই $\|\vec{v} - \vec{u}_1\| \geq \|\vec{v} - \vec{u}\|$ হবে। (*)-এর প্রমাণ খুবই সহজ--

$$\|\vec{v} - \vec{u}_1\|^2 = \left\| \underbrace{\vec{v} - \vec{u}}_{\vec{a}} + \underbrace{\vec{u} - \vec{u}_1}_{\vec{b}} \right\|^2 = \|\vec{a}\|^2 + \left\|\vec{b}\right\|^2 ,$$

যেহেতু $\left\langle \vec{a}, \vec{b} \right\rangle = 0$.

কোনো subspace-এর মধ্যে কোন্ vector-টা কোনো একটা vector-এর সবচেয়ে কাছে সেটা সবসময়েই এইভাবে orthogonal projection দিয়ে বার করা যায়। এই কাজটা অংকে তো বটেই, এমন কি engineering এবং statistics-এও খুব কাজে লাগে।

DAY 53
Adjoint

53.1 ছড়ায় ও ছবিতে

এবার আমরা এমন একটা জিনিস শিখব, যেটা শুনতে গুরুগম্ভীর হলেও আসলে আমরা সেটা হায়ার সেকণ্ডারীতেই শিখে থাকি। গুরুগম্ভীর নামটা হল **adjoint**. হায়ার সেকণ্ডারীতে যে নামে এটাকে চিনতে, সেটা হল transpose. অবশ্য adjoint আর transpose ঠিক একই জিনিস নয়, adjoint হল একটা general ধারণা, আর তার সবচেয়ে পরিচিত রূপটা হল transpose. প্রথমে অত খুঁটিনাটির মধ্যে না গিয়ে adjoint-কে transpose ধরে নিয়ে একটা অংক করি।

Example 20: Find the adjoint of the linear operator $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$T(x, y, z) = (x + y - z, x - z, y - z)$ for all $(x, y, z) \in \mathbb{R}^3$.[2] **(2020.CC12.3gii)**

SOLUTION: প্রথমে linear transformation-টাকে একটা matrix দিয়ে লিখে নাও--

$$T(x, y, z) = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} .$$

এবার এই matrix-টাকে transpose করে দিলে যে নতুন linear transformation-টা পাবে, সেটাই হবে T^*.

Here we are working with the canonical basis and standard inner product over field $\mathbb{R}$. So the matrix T^* is the transpose of the matrix of T. Hence

$$T^*(x, y, z) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y \\ x + z \\ -x - y - z \end{bmatrix} .$$

■

লক্ষ্য করো এখানে standard inner product, canonical basis ইত্যাদি গালভরা কিছু কথা লিখেছি। আসলে adjoint মানে তো সবসময়েই আর transpose নয়। যখন আমরা $\mathbb{R}$-এর উপরে standard inner product আর canonical basis নিয়ে কাজ করি, তখন সেটা transpose-এর রূপ ধরে। সেই জন্যেই ওটুকু বলে নিয়েছি। একই কায়দায় নীচের অংক দুটোও করো।

বুড়ো বলে উঠল, "এসব কী লিখেছ আবোল তাবোল? কাক বলল, "ও-সব লিখতে হয়। তা না হলে আদালতে হিসেব টিঁকবে কেন?"

--হ য ব র ল

Exercise 34: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear functional defined by $T(a, b) = (2a + b, a - 3b)$ and T^* be the adjoint of T. Then $T^*(3, 5)$ is equal to

(A) $(6, -5)$ (B) $(11, -12)$ (C) $(0, 0)$ (D) none of these.

(2021.CC12.1f)

HINT: এখানে ভাষার একটু সমস্যা আছে। যখন linear functional বলে, তার মানে একটা linear transformation যার codomain হল field-টা নিজেই[4] কিন্তু এখানে codomain-টা দেখা যাচ্ছে $\mathbb{R}^2$. তাই এখানে linear functional না বলে linear transformation বললেই সঠিক হত। ■

Exercise 35: Choose the correct answer with proper justification (1 mark for right answer and 1 mark for justification). If a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x, y, z) = (x + y - z, x - z, y - z)$, $\forall (x, y, z) \in \mathbb{R}^3$, then $T^*(x, y, z) =$

(A) $(x + y, x + z, -x - y - z)$

(B) $(x + y, x + z, y + z)$

(C) $(x + y, x + z, -x - y + z)$

(D) $(x + y, y + z, -x - y - z)$

(2022.CC12.1g) ■

এসব অংকে কাজ হচ্ছিল $\mathbb{R}$-এর উপরে। তাই transpose দিয়ে কাজ হল। যদি $\mathbb{C}$-এর উপরে কাজ হত, তবে conjugate transpose লাগত। যদি A একটা complex matrix হয়, তবে তার conjugate transpose হল

$$A^* = \overline{A'}.$$

এরকম একটা অংক দেখা যাক।

Example 21: If a linear transformation $T : \mathbb{C} \to \mathbb{C}^2$ is defined by $T(x) = (x, ix)$, for $x \in \mathbb{C}$, then find T^*.

SOLUTION: এখানে matrix-টা হল $\begin{bmatrix} 1 \\ i \end{bmatrix}$. এর conjugate transpose হল $\begin{bmatrix} 1 & -i \end{bmatrix}$. তাই $T^* : \mathbb{C}^2 \to \mathbb{C}$ হবে

$$T^*(x, y) = \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x - iy \text{ for } (x, y) \in \mathbb{C}^2.$$

[4]field-টা নিজেই নিজের উপরে একটা vector space হয়।

■

যদি matrix-টা real হয়, তবে field-টা $\mathbb{C}$ হলেও transpose-এই কাজ চলে যায়, কারণ A-র সবগুলো entry-ই যদি real হয়, তবে $A^* = A'$. নীচের অংকে তাই transpose-এর বেশী কিছু লাগবে না।

Example 22: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). Let T be a linear operator on $\mathbb{C}^3$ defined by $T(x, y, z) = (2x, 0, x)$, then the adjoint operator T^* of T is

(A) $T^*(x, y, z) = (2x + z, 0, 0)$

(B) $T^*(x, y, z) = (2x, 0, z)$

(C) $T^*(x, y, z) = (2x, 2y, 2z)$

(D) $T^*(x, y, z) = (0, 0, 2x + z)$.

(2020.CC12.1h)

SOLUTION:প্রথমে linear transformation-টাকে matrix দিয়ে লিখে নাও।

Here $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ 0 \\ x \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

এবার ওই A-টাকে transpose করে দিলে যে linear transformation-টা পাবে, সেটাই হবে T^*.

So

$$T^*\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A'\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + z \\ 0 \\ 0 \end{bmatrix}.$$

So the correct option is (A).

■

53.2 খুঁটিনাটি

এবার adjoint-এর খুঁটিনাটির প্রসঙ্গে আসি। এর জন্য শুরু করতে হবে দুটো inner product space নিয়ে, যাদের field হবে $\mathbb{R}$ বা $\mathbb{C}$ (দুজনেরই একই field চাই)। ধরো vector space দুটোর নাম দিলাম U আর V. এদের inner product দুটো ধরো $\langle \cdot, \cdot \rangle_U$ আর $\langle \cdot, \cdot \rangle_V$. এবার যদি তোমাকে কোনো linear transformation দিই, $T : U \to V$, তবে দেখানো যায় যে, এমন ঠিক একটাই linear transformation সম্ভব, $T^* : V \to U$, যাতে

$$\forall \vec{u} \in U \quad \forall \vec{v} \in V \quad \langle \vec{v}, T(\vec{u}) \rangle_V = \langle T^*(\vec{v}), \vec{u} \rangle_U$$

হয়। এই T^*-কেই বলে ওই inner product-এর সাপেক্ষে T-এর **adjoint**. এটা একটা theorem-এর আকারে লিখে প্রমাণ করে রাখা যাক।

THEOREM

Let U, V be finite dimensional inner product spaces over the same field. Let $T : U \to V$ be a linear transformation. Then there exists a unique linear transformation $T^* : V \to U$ such that

$$\forall \vec{u} \in U \quad \forall \vec{v} \in V \quad \langle \vec{v}, T(\vec{u})\rangle_V = \langle T^*(\vec{v}), \vec{u}\rangle_U \,,$$

where $\langle \cdot, \cdot \rangle_U$ and $\langle \cdot, \cdot \rangle_V$ are the inner products of U and V.

কোনো কোনো বইতে হয়তো দেখবে inner product দুটোকে উল্টো করে লেখা হয়েছে, মানে লিখেছে

$$\langle T(\vec{u}), \vec{v}\rangle_V = \langle \vec{u}, T^*(\vec{v})\rangle_U \,.$$

তাতে মানে কিছু বদলাবে না, কারণ inner product-এর ভিতরের জিনিস দুটোকে ওল্টালে খালি conjugate হয়ে যায় মাত্র।

Proof: এই প্রমাণে একটা জিনিস বারবারই কাজে লাগবে। সেটা বলে শুরু করি। যদি কোনো একটা inner product space-এর একটা ONB দেওয়া থাকে $\{\vec{b}_1, ..., \vec{b}_n\}$ তবে সেই inner product space-এর মধ্যে যেকোনো $\vec{v}$-কেই সেই ONB-র সাপেক্ষে এইভাবে লেখা যায়--

$$\vec{v} = \left\langle \vec{v}, \vec{b}_1 \right\rangle \vec{b}_1 + \cdots + \left\langle \vec{v}, \vec{b}_n \right\rangle \vec{b}_n.$$

এইটা মাথায় রেখে প্রমাণটা পড়।

<u>Existence</u>: Shall show

$\exists$linear transformation $T^* : V \to U \quad \forall \vec{u} \in U \quad \forall \vec{v} \in V \quad \langle \vec{v}, T(\vec{u})\rangle_V = \langle T^*(\vec{v}), \vec{u}\rangle_U \,.$

যেহেতু শুরুতেই একটা $\exists$ আছে, তাই প্রমাণ আরম্ভ হবে একটা উপযুক্ত T^* বানিয়ে। তার জন্য একটা লম্বা রাফ লাগবে। এখানে আমরা basis ব্যবহার করে linear transformation তৈরী করব। ধরো V-এর যা খুশি একটা basis নিলাম $\{\vec{v}_1, ..., \vec{v}_n\}$ আর U-এর যা খুশি n-খানা vector নিলাম, $\vec{w}_1, ..., \vec{w}_n$ (সবগুলো আলাদা আলাদা হবারও দরকার নেই)। তবে V থেকে U-তে ঠিক একটাই linear transformation আছে, যেটা $\vec{v}_i$-দের ওই $\vec{w}_i$-তে নিয়ে ফেলে। এটা আমরা এই বইতে আগেই দেখেছি। এবার কাজে লাগাব।
এর জন্য প্রথমে U আর V দুজনেরই basis নিয়ে শুরু করি।

Let $\{\vec{u}_1, ..., \vec{u}_m\}$ and $\{\vec{v}_1, ..., \vec{v}_n\}$ be ONBs of U, V, respectively.

তাহলে আমাদের দরকার--

$$\forall i, j \quad \langle \vec{v}_j, T(\vec{u}_i)\rangle_V = \langle T^*(\vec{v}_j), \vec{u}_i\rangle_U \,.$$

এটা বস্তুতঃ যেটা প্রমাণ করতে হবে, সেটারই অংশ। খালি যেকোনো $\vec{u} \in U$ আর যেকোনো $\vec{v} \in V$-এর জায়গায় $\vec{u}_i$ আর $\vec{v}_j$ নিয়েছি। আর এ থেকে বোঝার চেষ্টা করছি, $T^*(\vec{v}_j)$ কীভাবে নেওয়া উচিত। যেহেতু $T^*(\vec{v}_j) \in U$ এবং আমাদের হাতে U-এর একটা ONB আছে, তাই $T^*(\vec{v}_j) = \sum_{i=1}^m \langle T^*(\vec{v}_j), \vec{u}_i\rangle \vec{u}_i$ হবে।
এর মধ্যে $\langle T^*(\vec{v}_j), \vec{u}_i\rangle$ আমাদের পরিচিত জিনিস--

$$\langle T^*(\vec{v}_j), \vec{u}_i\rangle_U = \langle \vec{v}_j, T(\vec{u}_i)\rangle_V \,.$$

এখানে ডানদিকের জিনিসটায় কোনো T^* নেই। তাই ওটা ব্যবহার করে T^*-এর সংজ্ঞা বানানো যাবে। প্রথমে ডানদিকের জিনিসপত্রর নাম দিয়ে নিই--

Let $a_{ij} = \langle \vec{v}_j, T(\vec{u}_i)\rangle_V$ and $\vec{w}_j = \sum_i a_{ij}\vec{u}_i$.

তবে আমরা চাইছি যেন $\forall j \quad T^*(\vec{v}_j) = \vec{w}_j$ হয়। এইটা ভালো করে পড়ে বুঝে নাও।

Since $\{\vec{v}_1, ..., \vec{v}_n\}$ is a basis, hence by standard result, we know that there exists a (unique) linear transformation $T^* : V \to U$ such that $\forall j \quad T^*(\vec{v}_j) = \vec{w}_j$.

$\exists T^*$ We choose this T^*.

প্রমাণের প্রথম এবং সবচেয়ে কঠিন ধাপটা শেষ হল। এরপর দুটো $\forall$ আছে--

$\forall \vec{u}, \vec{v}$ Take any $\vec{u} \in U$ and $\vec{v} \in V$.

আমাদের দেখাতে হবে $\langle \vec{v}, T(\vec{u})\rangle_V = \langle T^*(\vec{v}), \vec{u}\rangle_U$. আমরা যেটুকু পেয়ে গেছি, সেটা হল--

By construction of T^*, we have

$$\forall i, j \quad \langle \vec{v}_j, T(\vec{u}_i)\rangle_V = \langle T^*(\vec{v}_j), \vec{u}_i\rangle_U . \qquad (*)$$

এবার আমরা $\vec{u}$ আর $\vec{v}$-কে $\vec{u}_i$ আর $\vec{v}_j$ দিয়ে প্রকাশ করে নিই--

Then $\vec{u} = \sum_1^m c_i\vec{u}_i$ and $\vec{v} = \sum_1^n d_j\vec{v}_j$ for some scalars c_i's and d_j's.
So

$$\begin{aligned}
\langle \vec{v}, T(\vec{u})\rangle_V &= \left\langle \sum_{j=1}^n d_j\vec{v}_j, T\left(\sum_{i=1}^m c_i\vec{u}_i\right)\right\rangle_V \\
&= \left\langle \sum_{j=1}^n d_j\vec{v}_j, \sum_{i=1}^m c_i T(\vec{u}_i)\right\rangle_V \quad [\because T \text{ is linear}] \\
&= \sum_{i=1}^m \sum_{j=1}^n \bar{c}_i d_j \langle \vec{v}_j, T(\vec{u}_i)\rangle_V
\end{aligned}$$

লক্ষ্য করো, c_i-গুলো কেমন $\bar{c}_i$ হয়ে গেল।

$$\begin{aligned}
&= \sum_{i=1}^m \sum_{j=1}^n \bar{c}_i d_j \langle T^*(\vec{v}_j), \vec{u}_i\rangle_U \quad [\text{ by } (*)] \\
&= \left\langle \sum_{j=1}^n d_j T^*(\vec{v}_j), \sum_{i=1}^m c_i\vec{u}_i\right\rangle_U \\
&= \left\langle T^*\left(\sum_{j=1}^n d_j\vec{v}_j\right), \sum_{i=1}^m c_i\vec{u}_i\right\rangle_U \quad [\because T^* \text{ is linear}] \\
&= \langle T^*(\vec{v}), \vec{u}\rangle_U ,
\end{aligned}$$

as required.

এবার unique দেখাব।

Uniqueness: Let $S : V \to U$ be any linear transformation such that

$$\forall \vec{u} \in U \quad \forall \vec{v} \in V \quad \langle \vec{v}, T(\vec{u}) \rangle_V = \langle S(\vec{v}), \vec{u} \rangle_U . \qquad (**)$$

Shall show that $S \equiv T^*$.
Since S, T^* are both linear transformations, it is enough to show that

$$\forall j \quad S(\vec{v}_j) = T^*(\vec{v}_j).$$

কারণ দুটো linear transformation যদি একটা basis-এর উপর মিলে যায়, তবে ওরা সর্বত্রই মিলে যেতে বাধ্য।

Take any j.
Then

$$\begin{aligned} S(\vec{v}_j) &= \sum_{i=1}^m \langle S(\vec{v}_j), \vec{u}_i \rangle \, \vec{u}_i \\ &= \sum_{i=1}^m \langle \vec{v}_j, T(\vec{u}_i) \rangle \, \vec{u}_i \quad [\text{ by } (**)] \\ &= \sum_{i=1}^m \langle T^*(\vec{v}_j), \vec{u}_i \rangle \, \vec{u}_i \\ &= T^*(\vec{v}_j), \end{aligned}$$

as required.

[Q.E.D]

এখানে বলা দরকার যে, এই theorem-টা শুধু finite dimensional vector space-দের জন্যই নয়, বিশেষ কিছু infinite dimensional vector space-এর জন্যেও খাটে। অবশ্য সে আলোচনা আমাদের এক্তিয়ারের বাইরে।

53.3 (Conjugate) transpose-এর সঙ্গে সম্পর্ক

আমরা একটু আগে বললাম যে, adjoint মানে $\mathbb{R}$-এর বেলায় transpose আর $\mathbb{C}$-এর বেলায় conjugate transpose. এক্ষুণি যে দাঁত-বার-করা theorem-টা দেখলাম, তা থেকে সে কথাটা বোঝা সহজ নয়। সেটা আলাদা একটা theorem-এর আকারে লিখে প্রমাণ করা যাক।

THEOREM

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let $T^* : V \to U$ be its adjoint wrt the standard inner products. If the matrix of T (wrt the canonical bases) is A, then the matrix of T^* (wrt the canonical bases) is A'.

Proof: আমাদের দেখাতে হবে--

$$\forall \vec{u} \in \mathbb{R}^m \quad \forall \vec{v} \in \mathbb{R}^n \quad \langle \vec{v}, T(\vec{u}) \rangle_V = \langle T^*(\vec{v}), \vec{u} \rangle_U ,$$

অর্থাৎ--

$$\forall \vec{u} \in \mathbb{R}^m \quad \forall \vec{v} \in \mathbb{R}^n \quad (A\vec{u})'\vec{v} = \vec{u}' A' \vec{v}.$$

এটা কি বলাই বাহুল্য নয়? **[Q.E.D]**

একইভাবে এই theorem-এর complex সংস্করণটাও প্রমাণ করে ফেলো--

Exercise 36: Let $T : \mathbb{C}^n \to \mathbb{C}^m$ be a linear transformation and let $T^* : V \to U$ be its adjoint (wrt the standard inner products). If the matrix of T (wrt the canonical bases) is A, then show that the matrix of T^* (wrt the canonical bases) is A^*. ■

যদি আমরা standard inner product ছাড়া অন্য কোনো inner product নিয়েও কাজ করি, তাহলেও (conjugate) transpose-এর কায়দাটা লাগানো যায়, তবে সেক্ষেত্রে matrix-টা নিতে হয় ওই inner product-দের জন্য ONB-দের সাপেক্ষে। সেটা নিয়েই নীচের অংকটা।

Exercise 37: ধরো U, V হল দুটো finite dimensional inner product space, এবং $T : U \to V$ হল কোনো linear transformation. আমরা U-এর কোনো একটা ONB এবং V-এর কোনো একটা ONB ঠিক করে নিলাম, এবং তাদের সাপেক্ষে T এবং T^*-এর matrix বার করলাম। দেখাও যে, এ দুটো matrix পরস্পরের (conjugate) transpose হতে বাধ্য।
Hint:
এর মধ্যে হাতি-ঘোড়া কিছু নেই। যেই তুমি একটা ONB-র সাপেক্ষে একটা inner product space-কে লিখবে, অমনি inner product-টা standard inner product-এ পরিণত হবে। যেমন, ধরো $\{\vec{u}_1, ..., \vec{u}_n\}$ যদি U-এর একটা ONB নিলাম। এবার $\vec{u}, \vec{w} \in U$ নিয়ে $\langle \vec{u}, \vec{w} \rangle$ বার করতে চাই। তার জন্য ওদেরকে ONB-র সাপেক্ষে লিখি--

$$\vec{u} = a_1\vec{u}_1 + \cdots + a_n\vec{u}_n \text{ এবং } \vec{w} = b_1\vec{u}_1 + \cdots + b_n\vec{u}_n.$$

তবে অমনি দেখবে $\langle \vec{u}, \vec{w} \rangle = \sum a_i \bar{b}_i$ হচ্ছে। ■

নীচের অংক কয়েকটা সামান্য অন্য স্বাদের। এই অংকগুলোর জন্মরহস্য বোঝার জন্য একটা ব্যাপার লক্ষ্য করো। যদি $T : \mathbb{R}^n \to \mathbb{R}$ কোনো linear functional হয়, তবে $T(x_1, ..., x_n)$-এর চেহারা হবে $T(x_1, ..., x_n) = \sum a_i x_i$-এর মত। এটাকে আমরা লিখতে পারি $T(\vec{x}) = \vec{a}'\vec{x}$, যেটা $\vec{a}$ আর $\vec{x}$-এর মধ্যে standard inner product. যদি field-টা $\mathbb{C}$ হত, তবে a_i-এর বদলে $\overline{a_i}$ নিলেই একইভাবে $T(\vec{x}) = \vec{a}^*\vec{x}$ লেখা যেত, যেটা $\mathbb{C}$-এর বেলায় $\vec{a}$ আর $\vec{x}$-এর standard inner product. এই ব্যাপারটা আসলে যেকোনো finite dimensional inner product space-এর জন্যই হয়, এবং খানিকটা বাড়তি শর্ত চাপালে infinite dimension-এর ক্ষেত্রেও। সেটা নিয়েই নীচের অংকগুলো।

Exercise 38: Let U be a finite dimensional inner product space over $\mathbb{C}$. Let $T : U \to F$ be any linear functional. Show that there must exist some $\vec{u} \in U$ such that

$$\forall \vec{x} \in U \quad T(\vec{x}) = \langle \vec{x}, \vec{u} \rangle .$$

এই জিনিসটা অংকের জগতে একটা অত্যন্ত গুরুত্বপূর্ণ জিনিস। একে বলে **Riesz representation theorem.** এখানে আমরা finite dimensional inner product space নিয়ে কাজ করছি বলে প্রমাণটা সহজই। কিন্তু এর আসল শক্তি প্রকাশ পায় infinite dimension-এর বেলায়।
■

53.4 Adjoint বার করে কী লাভ?

এখানে স্বাভাবিকভাবেই প্রশ্ন ওঠে--adjoint নিয়ে মাথা ঘামাচ্ছি কেন? এর সংক্ষিপ্ত উত্তর হল infinite dimensional vector space-দের বেলায় transpose-এর ধারণাটাকে generalise করার জন্য। অবশ্য অমনি তুমি প্রশ্ন করে বসতে পারো--transpose-এর ধারণাটাকে generalise করে লাভ কী? সেইটা কিছু শক্ত প্রশ্ন। পুরো ব্যাখ্যা করে বোঝানো এই বইয়ের পাল্লার বাইরে। তবে একটা প্রয়োগ বলতে পারি। যখন আমরা square matrix-দের নিয়ে কাজ করি, তখন real symmetric matrix বা complex Hermitian matrix-দের অন্যতম সুবিধা হল ওদের eigenvalue-গুলো সবাই real হয়। কিন্তু infinite dimensional vector space-এর বেলায় তো আর real symmetric বা complex Hermitian বলতে পারছি না, কারণ ওখানে তো কোনো matrix-ই নেই। সেখানে আমরা adjoint ব্যবহার করে একই কাজ করতে পারি। Real symmetric বা complex Hermitian না বলে আমরা সেখানে বলি **self-adjoint**, অর্থাৎ এমন একটা linear transformation, যে নিজেই নিজের adjoint. এরকম linear transformation-দের বেলাতেও eigenvalue-রা সবাই real হয়।

53.5 Adjoint-এর কিছু ধর্ম

এবার আমরা adjoint-এর কিছু সহজ ধর্ম আলোচনা করব। প্রথম ধর্মটা হল যদি কোনো linear transformation নিয়ে তার adjoint-এর adjoint নাও, তবে ফের মূল linear transformation-এই ফিরে আসবে।

> **THEOREM**
>
> Let $T : U \to V$ be a linear transformation, where U, V are finite dimensional inner product spaces. Then
> $$(T^*)^* = T.$$

Proof: এর সবচেয়ে সহজ প্রমাণ হল matrix দিয়ে, কারণ দুবার conjugate transpose নিলে ফের শুরুর matrix-এই ফিরে আসতে হয়।

> Let us fix some ONBs of U and V.
> We know that the matrix representation of the adjoint of a linear transformation is the conjugate transpose of the matrix of the original linear transformation wrt the two bases.
> So if A is the matrix for T wrt these bases, then the matrix for T^* is A^* and that for $(T^*)^*$ is $(A^*)^* = A$.
> Since for a given pair of bases a linear transformation is completely determined by its matrix, hence $(T^*)^* = T$.

[Q.E.D]

ধরো দুটো linear transformation আছে $T_1, T_2 : U \to V$. যদি এদের কোনো linear combination নিই, যেমন ধরো $3T_1 - 4T_2$, তবে সেটা ফের U থেকে V-তে একটা linear transformation-ই হবে। যদি U, V দুজনেই inner product space হয়, তবে T_1, T_2 এবং $3T_1 - 4T_2$, এই তিনজনেরই adjoint নেওয়া যায়। এদের মধ্যে কি কোনো সম্পর্ক আছে? আছে, এবং সেটা অত্যন্ত সহজ, স্বাভাবিক--

THEOREM

Let $T_1, T_2 : U \to V$ be two linear transformations between the inner product spaces U and V. If a_1, a_2 are two scalars, then

$$(a_1T_1 + a_2T_2)^* = \overline{a}_1T_1^* + \overline{a}_2T_2^*.$$

এইটারও প্রমাণ হবে ONB-দের সাপেক্ষে matrix দিয়ে। প্রমাণটা গুছিয়ে লেখার দায়িত্ব তোমার উপরেই রইল।

53.6 Adjoint বনাম adjugate

এখানে notation নিয়ে একটা কথা বলা দরকার। আগেকার দিনে adjoint বলতে লোকে বুঝত cofactor-এর matrix-এর transpose, এবং A কোনো square matrix হলে তার সেই adjoint-কে লিখত $\text{adj}(A)$. এবং সেটা খালি square matrix-দের ক্ষেত্রেই ব্যবহৃত হত। কিন্তু আধুনিক মতে adjoint-এর নতুন সংজ্ঞা হবার পর cofactor-এর matrix-এর transpose-টার নাম দেওয়া হয়েছে **adjugate**. এখানেও প্রথম তিনটে অক্ষর একই থাকায়, লেখার সময়ে ওই $\text{adj}(A)$-ই লেখা হয়। যদি কোনো পুরোনো বই পড়, তবে হয়তো দেখবে $\text{adj}(A)$-কে adjoint বলা হচ্ছে, তখন গুলিয়ে ফেলো না যেন!

Answers

1. $\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 + a_4\vec{v}_4$, যেখানে $a_1 = \frac{4}{4} = 1, a_2 = \frac{-8}{4} = -2, a_3 = \frac{2}{4} = \frac{1}{2}, a_4 = \frac{6}{4} = \frac{3}{2}$. **2.** 0. **3.** $\vec{v} = \vec{0}$. **5.** ওটার জন্য সংজ্ঞার প্রথম শর্তটা কাজ করবে না। কারণ যদি $\vec{x} = (0, 1, 0, 0)$ নিই, তবে $\langle \vec{x}, \vec{x} \rangle < 0$ হয়ে যাবে। **6.**

<u>(If part):</u> Let $\forall i \quad a_i > 0$. Then $\langle \vec{x}, \vec{x} \rangle = \sum a_i x_i^2 \geq 0$. Also it is 0 if and only if $\vec{x} = \vec{0}$.

$\langle \vec{x}, \vec{y} \rangle = \sum a_i x_i y_i = \sum a_i y_i x_i = \langle \vec{y}, \vec{x} \rangle$.

For $\alpha, \beta \in \mathbb{R}$ we have $\langle \alpha\vec{x} + \beta\vec{y}, \vec{z} \rangle = \sum a_i(\alpha x_i + \beta y_i)z_i = \alpha \sum a_i x_i z_i + \beta \sum a_i y_i z_i = \alpha \langle \vec{x}, \vec{z} \rangle + \beta \langle \vec{y}, \vec{z} \rangle$.

<u>(Only if part):</u> Let $\langle \vec{x}, \vec{y} \rangle = \sum a_i x_i y_i$ be an inner product.

Let, if possible, $a_i \leq 0$ for some i.

Let $\vec{x} = \vec{e}_i \neq \vec{0}$.

Then $\langle \vec{x}, \vec{x} \rangle = a_i \leq 0$, which is not possible.

7. না, কারণ $f(\vec{e}_2, \vec{e}_2) = 0$. যেহেতু $\vec{e}_2 \neq \vec{0}$, তাই inner product হওয়ার জন্য $f(\vec{e}_2, \vec{e}_2) > 0$ হওয়া উচিত ছিল। **8.** না, inner product নয়, কারণ $\vec{\alpha} = (0, 1, -1) \neq \vec{0}$ হলেও $\langle \vec{\alpha}, \vec{\alpha} \rangle = 0$ হয়ে যাচ্ছে। **9.** না, কারণ আমরা $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle$ ব্যবহার করেছিলাম, যেটা $\mathbb{C}$-এর উপরে নাও খাটতে পারে। যেমন যদি $\alpha = (1, 1)$ এবং $\beta = (i, i)$ নিই, তবে $\|\alpha + \beta\| = \|\alpha - \beta\|$ হচ্ছে, কিন্তু $\langle \alpha, \beta \rangle$ মোটেই শূন্য হচ্ছে না।
11. $\|\alpha + \beta\|^2 = \langle \alpha + \beta, \alpha + \beta \rangle = \|\alpha\|^2 + \|\beta\|^2 + 2\langle \alpha, \beta \rangle = \|\alpha\|^2 + \|\beta\|^2$ হওয়া মানে $\langle \alpha, \beta \rangle = 0$.
13. না, orthogonal নয়, কারণ $\langle (1, 2, 3), (1, 1, 1) \rangle = 1 \times 1 + 2 \times 1 + 3 \times 1 \neq 0$. **14.** খালি দুটোই আছে $\left(\frac{4}{5}, -\frac{3}{5}\right)$ এবং $\left(-\frac{4}{5}, \frac{3}{5}\right)$. **16.** $\langle \vec{u}_2, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle - \langle \vec{u}_1, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle - \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \langle \vec{v}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle - \langle \vec{u}, \vec{v} \rangle = 0$.
17. অবশ্যই! **18.** $\|A\vec{v}\|^2 = (A\vec{v})^*(A\vec{v}) = \vec{v}^*A^*A\vec{v} = \vec{v}^*\vec{v} = \|\vec{v}\|^2$, যেহেতু $A^*A = I$. এ থেকেই পাচ্ছি $\|A\vec{v}\| = \|\vec{v}\|$, কারণ norm সর্বদা ≥ 0 হয়। **19.** $\|A\vec{e}_i\| = \|\vec{e}_i\| = 1$. তাই A-র i-th column-টার norm হচ্ছে 1. আবার $i \neq j$ হলে $\|A(\vec{e}_i + \vec{e}_j)\| = \|\vec{e}_i + \vec{e}_j\| = \sqrt{2}$. যদি A-র i-th column-কে $\vec{a}_i$ বলি, তবে পাচ্ছি

$\|\vec{a}_1 + \vec{a}_2\| = \sqrt{2}$. মানে $\|\vec{a}_1\|^2 + \|\vec{a}_2\|^2 + 2\langle \vec{a}_i, \vec{a}_j\rangle = 2$. সুতরাং $\langle \vec{a}_i, \vec{a}_j\rangle = 0$. **20.** বিজোড় হতে বাধ্য, কারণ সবগুলো eigenvalue-এর গুণফল (প্রত্যেকটাকে algebraic multiplicity বার ধরে) তো determinant-এর, মানে -1-এর, সমান হতে হবে। **21.** নইলে $\vec{v}_2 = \vec{0}$, মানে $\vec{u}_2 = \frac{\langle \vec{u}_2, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\vec{u}_1$ হত। কিন্তু তাহলে তো $\vec{u}_2$ হয়ে যাচ্ছে $\vec{u}_1$-এর একটা multiple! কিন্তু সেটা অসম্ভব কারণ $\vec{u}_1, \vec{u}_2$ বলা ছিল linearly independent. **22.** প্রথমে সবকিছুর নাম দিয়ে নিই--

> Let
>
> $$\begin{aligned}\vec{u}_1 &= (1,0,1),\\ \vec{u}_2 &= (1,0,-1),\\ \vec{u}_3 &= (1,3,4).\end{aligned}$$

তারপর orthogonal বানাই--

> Gram-Schmidt process:
>
> $$\begin{aligned}\vec{v}_1 &= \vec{u}_1 = (1,0,1)\\ \langle \vec{v}_1, \vec{v}_1\rangle &= 2,\\ \langle \vec{u}_2, \vec{v}_1\rangle &= 0,\\ \vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\vec{v}_1\\ &= (1,0,-1)\\ \langle \vec{v}_2, \vec{v}_2\rangle &= 2,\\ \langle \vec{u}_3, \vec{v}_1\rangle &= 5,\\ \langle \vec{u}_3, \vec{v}_2\rangle &= -3,\\ \vec{v}_3 &= \vec{u}_3 - \frac{\langle \vec{u}_3, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\vec{v}_1 - \frac{\langle \vec{u}_3, \vec{v}_2\rangle}{\langle \vec{v}_2, \vec{v}_2\rangle}\vec{v}_2\\ &= (0,3,0).\end{aligned}$$

এবার norm দিয়ে ভাগ করার পালা, যাকে বলে normalisation–

> <u>Normalisation:</u>
>
> $$\begin{aligned}\vec{w}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} = \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)\\ \vec{w}_2 &= \frac{\vec{v}_2}{\|\vec{v}_2\|} = \left(\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}\right)\\ \vec{w}_3 &= \frac{\vec{v}_3}{\|\vec{v}_3\|} = (0,1,0).\end{aligned}$$
>
> Thus a required ONB for $\mathbb{R}^3$ is $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$.

23.

Let

$$\begin{aligned} \vec{u}_1 &= (1,1,0,1), \\ \vec{u}_2 &= (1,-2,0,1), \\ \vec{u}_3 &= (1,0,-1,2). \end{aligned}$$

Gram-Schmidt orthogonalisation:

$$\begin{aligned} \vec{v}_1 &= \vec{u}_1 = (1,1,0,1) \\ \langle \vec{v}_1, \vec{v}_1 \rangle &= 3, \\ \langle \vec{u}_2, \vec{v}_1 \rangle &= -1, \\ \vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle}\vec{v}_1 \\ &= \left(\tfrac{4}{3}, -\tfrac{5}{3}, 0, \tfrac{1}{3}\right) \\ \langle \vec{v}_2, \vec{v}_2 \rangle &= \tfrac{14}{3}, \\ \langle \vec{u}_3, \vec{v}_1 \rangle &= 3, \\ \langle \vec{u}_3, \vec{v}_2 \rangle &= 2, \\ \vec{v}_3 &= \vec{u}_3 - \frac{\langle \vec{u}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle}\vec{v}_1 - \frac{\langle \vec{u}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle}\vec{v}_2 \\ &= \left(-\tfrac{4}{7}, -\tfrac{2}{7}, -1, \tfrac{6}{7}\right) \end{aligned}$$

Normalisation:

$$\begin{aligned} \vec{w}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} = \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, 0, \tfrac{1}{\sqrt{3}}\right) \\ \vec{w}_2 &= \frac{\vec{v}_2}{\|\vec{v}_2\|} = \tfrac{1}{\sqrt{42}}\,(4,-5,0,1) \\ \vec{w}_3 &= \frac{\vec{v}_3}{\|\vec{v}_3\|} = \tfrac{1}{\sqrt{105}}\,(-4,-2,-7,6) \end{aligned}$$

24. $\frac{1}{\sqrt{2}}(1,0,1),\ (0,1,0),\ \frac{1}{\sqrt{2}}(-1,0,1).$ **25.** $\left\{\frac{1}{\sqrt{3}}(1,1,1), \frac{1}{\sqrt{78}}(5,-7,2), \frac{1}{\sqrt{26}}(3,1,-4)\right\}.$

26. $\left\{\frac{1}{\sqrt{3}}(1,1,0,1),\ \ \frac{1}{\sqrt{6}}(1,1,0,-2),\ \ \frac{1}{\sqrt{2}}(1,-1,0,0)\right\}.$ **27.**

Putting $\vec{v} = \vec{v}_k$ in the given condition, we have $a_i = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$.

So $\|\vec{v}_k\|^2 = \sum_i |a_i|^2 = 1.$

Next, putting $\vec{v} = \vec{v}_k + \vec{v}_\ell$ for some for $k \neq \ell$,

we have $a_i = \begin{cases} 1 & \text{if } i = k, \ell \\ 0 & \text{otherwise} \end{cases}$.

So $\|\vec{v}_k + \vec{v}_\ell\|^2 = \sum_i |a_i|^2 = 2.$
Thus, $\|\vec{v}_k\|^2 + \|\vec{v}_\ell\|^2 + 2\langle \vec{v}_k, \vec{v}_\ell\rangle = 2.$
Since $\|\vec{v}_k\| = \|\vec{v}_\ell\| = 1,$ this means $\langle \vec{v}_k, \vec{v}_\ell\rangle = 0.$

29. $\text{span}\{(1,-1,0),\,(1,0,-1)\}$.
30. $W^\perp = \{a + bt + ct^2 + dt^3 \ :\ a + \frac{b}{2} + \frac{c}{3} + \frac{d}{4} = 0,\, \frac{a}{2} + \frac{b}{3} + \frac{c}{4} + \frac{d}{5} = 0\}$. এর একটা basis হল $\left\{1 - 6t + 6t^2,\, 2 - 9t + 10t^3\right\}$. **31.** $U \cap W$ যেহেতু একটা subspace, তাই $\vec{0} \in U \cap W$. এবার ধরো $\vec{v} \in U \cap W$. তাহলে $\vec{v} \in U$-ও বটে আবার $\vec{v} \in W$-ও বটে। সুতরাং $\langle \vec{v}, \vec{v}\rangle = 0,$ মানে $\vec{v} = \vec{0}$. **32.** (B) আর (D) খালি ঠিক। **33.** $\vec{0}$ সবার সঙ্গেই orthogonal. তাই $\vec{0} \in W$. এবার ধরো $\vec{v}, \vec{w} \in W$ আর a, b দুটো scalar. তাহলে যেকোনো $\vec{u} \in S$-এর জন্যই $\langle a\vec{v} + b\vec{w}, \vec{u}\rangle = a\langle a\vec{v}, \vec{u}\rangle + b\langle \vec{w}, \vec{u}\rangle = 0.$ সুতরাং $a\vec{v} + b\vec{w} \in W.$ সুতরাং W হচ্ছে closed under linear combination. তাই W একটা subspace. **34.** (B). এখানে $T^* = T$ হচ্ছে। **35.** (A).
38. U-এর একটা ONB নিই, $\{\vec{v}_1, ..., \vec{v}_n\}$. এবার $u_i = \overline{T(\vec{v}_i)}$ নাও, এবং তাদের ব্যবহার করে $\vec{u}$ বানাও $\vec{u} = u_1\vec{v}_1 + \cdots u_n\vec{v}_n$. এবার যখন কোনো $\vec{x} \in U$ নেবে, সেটাকে আমরা $\vec{x} = \sum x_i\vec{v}_i$ আকারে লিখে নেব। তবে $T(\vec{x}) = \sum x_i T(\vec{v}_i) = \sum x_i \overline{u_i}$ হবে। এদিকে $\langle \vec{x}, \vec{u}\rangle$-ও $\sum x_i \overline{u_i}$-এর সমান।

Chapter IX
Spectral decomposition

DAY 61
Orthogonal/unitary similarity

এ বইতে ইতিমধ্যেই আমরা দু রকমের equivalence নিয়ে আলোচনা করেছি, similarity এবং congruence. এবার আরেকটা নতুন ধরণের equivalence-এর কথা শিখব--

- field-টা $\mathbb{R}$ হলে একে বলে **orthogonal similarity**,
- আর field-টা $\mathbb{C}$ হলে একে বলে **unitary similarity**.

এই অধ্যায়ের আলোচনাটা খালি সেইসব ক্ষেত্রের জন্যই, যেখানে field-টা $\mathbb{R}$ বা $\mathbb{C}$.

DEFINITION: Orthogonal similarity

Two square matrices, A and B, over $\mathbb{R}$ are called **orthogonally similar** if there is an orthogonal matrix P such that $A = P'BP$.

লক্ষ্য করো, যেহেতু P-কে orthogonal নিয়েছি, তাই $P' = P^{-1}$-ও বটে। তাই আমরা চাইলে $A = P^{-1}BP$-ও লিখতে পারতাম। তার মানে orthogonally similar হলে congruent-ও হয়, আবার similar-ও হয়।

DEFINITION: Unitary similarity

Two square matrices, A and B, over $\mathbb{C}$ are called **unitarily similar** if there is a unitary matrix P such that $A = P^*BP$.

61.1 Basis দিয়ে ভাবা

আমরা যখন similarity বা congruence আলোচনা করেছিলাম, তখন ওদের পিছনে একটা basis পরিবর্তনের গল্প ছিল। যেমন similarity-র বেলায় আমরা কোনো একটা linear transformation নিয়ে শুরু করেছিলাম, $f : V \to V$. আমাদের উদ্দেশ্য ছিল সেটাকে একটা matrix দিয়ে লেখা। কিন্তু তার জন্য একটা basis নিতে হচ্ছিল। যদি basis-টা বদলায়, তবে matrix-টাও বদলে যেতে পারে। এই ব্যাপার থেকেই similarity-র উদ্ভব--দুটো matrix-কে similarity বলতাম, যদি ওরা একই linear transformation-কে প্রকাশ করে (বিভিন্ন basis-এর সাপেক্ষে)। একইভাবে congruence-এর ধারণাটা

এসেছিল একই quadratic form-কে বিভিন্ন basis-এর সাপেক্ষে প্রকাশ করতে গিয়ে। স্বভাবতঃই প্রশ্ন জাগে orthogonal similarity বা unitary similarity-র পিছনে এরকম কোনো basis-এর গল্প আছে কিনা। উত্তর হল--হ্যাঁ, আছে। এবার সেটাই আলোচনা করব।

এখানে আমরা কাজ করব inner product space নিয়ে। ধরো V একটা inner product space. আমরা field-টা নেব $\mathbb{R}$. ধরো $f : V \to V$ হল কোনো একটা linear transformation. এবার আমরা V-এর যা খুশি একটা ONB (মানে orthonormal basis) নেব। লক্ষ্য করো, আমরা ONB নিয়েছি, যেকোনো সাধারণ basis নয়। এর সাপেক্ষে ধরো f-এর matrix-টা হল A. এবার আমরা আরেকটা ONB নেব। এর জন্য f-এর matrix-টা হল ধরো B. তবে A আর B পরস্পরের সঙ্গে orthogonally similar হবে। বিপরীতপক্ষে যদি A আর B যেকোনো দুটো orthogonally similar matrix হয়, তবে ওরা একই linear transformation-কে দুটো ONB-র সাপেক্ষে প্রকাশ করতে বাধ্য।

এখানে আমরা আলোচনাটা করলাম linear transformation নিয়ে। একই কাজ করতে পারতাম quadratic form নিয়েও। ধরো $q : V \times V \to \mathbb{R}$ হল কোনো একটা quadratic form. ফের দুটো ONB নাও। যদি ওদের সাপেক্ষে q-এর matrix দুটো হয় A আর B (দুজনেই real, symmetric matrix), তবে ওরা পরস্পরের সঙ্গে orthogonally similar হতে বাধ্য। বিপরীতপক্ষে যদি A আর B দুটো real, symmetric matrix হয়, যারা পরস্পরের সঙ্গে orthogonally similar, তবে ওরা বিভিন্ন ONB-র সাপেক্ষে একই quadratic form-কে প্রকাশ করবে।

এই দুটো জিনিসই প্রমাণ করা খুব সহজ। খালি এটা লক্ষ্য করলেই হবে যে, দুটো ONB-র মধ্যে change of basis matrix-টা সর্বদা orthogonal হয়। একটা ONB-কে বদলে আরেকটা ONB তৈরী করাটা ছবি দিয়ে ভাবতে সুবিধা। যেহেতু basis-এর vector-গুলোর দৈর্ঘ্য বা ওদের মধ্যবর্তী কোণগুলো অপরিবর্তিত থাকছে, তাই এখানে খালি rotation (ঘূর্ণন) এবং reflection (প্রতিফলন, মানে সহজভাষায় ওল্টানো) ছাড়া আর কিছুই হচ্ছে না।

এতক্ষণ আমরা খালি $\mathbb{R}$-এর উপর কাজ করছিলাম। যদি $\mathbb{C}$-এর উপর কাজ করো, তবে একইভাবে unitary similarity পাবে।

এবার আমরা unitary similarity এবং orthogonal similarity-র দুটো গুরুত্বপূর্ণ প্রয়োগ আলোচনা করব। প্রথমটা হল Schur decomposition, যার সাথে আগেই আমাদের পরিচয় হয়েছিল। তখন প্রমাণটা দিই নি। এবার দেব। আর দ্বিতীয় প্রয়োগটা হল linear algebra-র একটা অন্যতম বিখ্যাত theorem, যার নাম spectral decomposition.

61.2 Schur decomposition-এর প্রমাণ

সহজ ভাষায় Schur decomposition মানে হল--যেকোনো square matrix-ই কোনো না কোনো upper triangular matrix-এর সঙ্গে unitarily similar. এখানে field-টা $\mathbb{C}$. সাবধান, $\mathbb{R}$-এর বেলায় কিন্তু এই theorem-টা কাজ করবে না। আরেকটা কথা, এখানে upper triangular-এর মধ্যে কোনো আলাদা জাদুমন্ত্র নেই, একই জিনিস করা যায় lower triangular নিয়েও।

Schur decomposition

Any $A_{n\times n}$ over $\mathbb{C}$ can be written as

$$A = PTP^*$$

for some unitary P and some (upper) triangular T.

Proof: প্রমাণটা হবে n-এর উপর induction লাগিয়ে।

> We shall show that for any $n \times n$ matrix A there is a unitary matrix P such that P^*AP is upper triangular.
>
> We shall use induction on n.

Induction মানেই পরিচিত সেই তিনটে অংশ--basis, hypothesis আর step--

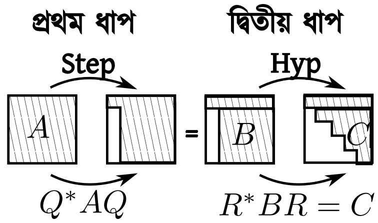


Fig 1

> Basis: If $n = 1$, then A is just a number, and so is already an upper triangular matrix. Hence we can take $P = 1$.
>
> Hypothesis: Let us assume the result for $n = 1, ..., m$.

এবারই আসল কাজ--

> Step: Shall prove the result for $n = m + 1$.

এবার কী করতে চলেছি সেটা বোঝার জন্য Fig 1-এর দিকে চোখ রাখো। কাজটা দুই ধাপে হবে। প্রথম ধাপে আমরা unitary similarity দিয়ে খালি প্রথম column-টার ব্যবস্থা করব। এমন একটা unitary matrix বার করব, Q, যাতে Q^*AQ-এর প্রথম column-এর শেষের $n-1$-খানা entry শূন্য হয় (ছবিতে matrix-এর ভিতরের সাদা অংশগুলোতে কেবল শূন্য আছে)। এর ফলে যে matrix-টা পাব, তার নীচের ডানদিকের $(n-1)\times(n-1)$ সাইজের submatrix-টা নেব (ছবিতে যেটাকে B নাম দিয়েছি)। এইটা যেহেতু সাইজে ছোটো, তাই এর উপরে induction hypothesis লাগানো যাবে। তাই আমরা এমন একটা unitary matrix পাব (ধরো $R_{(n-1)\times(n-1)}$), যাতে $R^*BR = C$ হবে upper triangular. ছবির শেষ matrix-টাকে দেখলেই বুঝবে যে, সব মিলিয়ে আমরা একটা upper triangular matrix পেয়ে গেছি। এবার Q এবং R-কে মিলিয়ে P তৈরী করে ফেললেই হবে। সেটা আমরা এক্ষুণি আলোচনা করব। আগে বলি Q কী করে বার করতে হয়।
আমরা চাইছি যেন Q^*AQ-এর প্রথম column-এর নীচের $n-1$-খানা entry শূন্য হয়, মানে প্রথম column-টা $\lambda\vec{e}_1$ চেহারার হয়--

$$Q^*AQ\text{-এর প্রথম column} = Q^*A \times (Q\text{-এর প্রথম column}) = \lambda\vec{e}_1,$$

অর্থাৎ, যেহেতু Q হল unitary, তাই Q^*-কে Q হিসেবে ডানদিকে নিয়ে--

$$A \times (Q\text{-এর প্রথম column}) = \lambda Q\vec{e}_1 = \lambda \times (Q\text{-এর প্রথম column}),$$

সুতরাং দেখতে পাচ্ছ যে, Q এমন একটা unitary matrix, যার প্রথম column হল A-র একটা eigenvector.

> Let $\vec{v} \neq \vec{0}$ be any eigenvector of A (for some eigenvalue λ).
>
> Let $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$.

Unit vector বানিয়ে নিলাম, নইলে unitary matrix-এর column হতে পারবে না।

> We extend $\vec{u}$ to an ONB of $\mathbb{C}^n$ and let $Q_{n\times n}$ be the unitary matrix with these vectors as its columns.

এখানে কী করলাম বুঝলে তো? ONB-র vector-গুলোকে পাশাপাশি column-এর মত দাঁড় করিয়ে একটা matrix বানালাম Q. যেহেতু column-গুলো মিলে একটা ONB হয়, তাই Q-টা unitary হতে বাধ্য। ব্যস্, প্রথম ধাপ শেষ! গুছিয়ে লিখে দিই, এই ধাপের ফলে প্রথম column-টার কী চেহারা হয়েছে--

> Then the first column of Q^*AQ is
>
> $$Q^*A\vec{u} = Q^*(\lambda\vec{u}) = \lambda Q^*\vec{u} = \begin{bmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

এবার দ্বিতীয় ধাপের শুরুতে Q^*AQ-র নীচের $(n-1)\times(n-1)$ অংশটা বার করে নিই।

> So we can partition Q^*AQ as
>
> $$Q^*AQ = \begin{bmatrix} \lambda & L \\ \vec{0} & B \end{bmatrix},$$
>
> for some $1\times(n-1)$ matrix L and some $(n-1)\times(n-1)$ matrix B.
> By induction hypothesis there is some unitary matrix $R_{(n-1)\times(n-1)}$ such that $R^*BR = C$ is upper triangular.

ব্যস্, আমাদের হাতে এখন $Q_{n\times n}$ এবং $R_{(n-1)\times(n-1)}$ এসে গেছে। এদের মিলিয়ে আমরা P বানাব। মনে রেখো R দিয়ে আমরা পুরো matrix-টাকে গুণ করি নি (R সাইজে ছোটো), গুণ করেছি কেবল নীচের ডানদিকের $(n-1)\times(n-1)$ সাইজের submatrix-টাকে। এবার আমরা আমরা একটা কৌশল করব--

> Let $S_{n\times n} = \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R \end{bmatrix}$

এই S দিয়ে পুরো matrix-টাকে গুণ করলেই নীচের ডানদিকের submatrix-টাকে R দিয়ে গুণ করা হয়! এই S-টা আবার unitary-ও বটে--

> Then S is unitary.
>
> [[<u>Because:</u>
>
> S is square and
>
> $$\begin{aligned} S^*S &= \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R^* \end{bmatrix}\begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R \end{bmatrix} = \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R^*R \end{bmatrix} \\ &= \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & I_{n-1} \end{bmatrix} = I_n. \end{aligned}$$

]]

We take $P = QS$. Since product of unitary matrices is unitary, so P is unitary.

Also

$$P^*AP = S^*(Q^*AQ)S = \text{upper triangular},$$

as required.

So the result is proved by the principle of mathematical induction.

[Q.E.D]

Exercise 1: ধরো A একটা complex, square matrix, যার Schur decomposition হল $P^*AP = T$. তবে দেখাও যে, T-এর diagonal entry-গুলো হল A-র eigenvalue. ■

61.3 Spectral decomposition

আমরা দেখলাম যেকোনো square matrix-ই কোনো না কোনো (upper) triangular matrix-এর সঙ্গে similar. আরও দেখলাম যে সব square matrix মোটেই semisimple নয়। একটা square matrix দেওয়া থাকলে সেটা semisimple কিনা সেটা পরীক্ষা করে দেখার একটা কায়দা হল যাবতীয় eigenvalue আর eigenvector বার করে দেখা কোথাও geometric multiplicity-টা algebraic multiplicity-র থেকে ছোটো হয়ে যাচ্ছে কিনা। বলাই বাহুল্য যে এইভাবে খড়ের গাদায় সূঁচ খোঁজার মত করে খোঁজা খুবই কঠিন কাজ। তবে এক ধরণের matrix আছে যারা অতি অবশ্যই semisimple হয়, তারা হল real, symmetric matrix. মানে তুমি যাই real, symmetric matrix A নাও না কেন সেটাকে $A = PDP^{-1}$ আকারে লেখা যাবে, যেখানে D একটা diagonal matrix আর P একটা nonsingular matrix. শুধু তাই নয়, এর সাথে একটা বোনাসও আছে--চাইলে P-টাকে তুমি orthogonal-ও নিতে পারো। তাহলে $A = PDP^{-1} = PDP'$ হবে। এইটা একটা অতি বিখ্যাত theorem, একে বলে spectral decomposition.

Spectral decomposition

Any real, symmetric matrix A can be written as

$$A = PDP'$$

for some real, orthogonal matrix P and some real, diagonal matrix D.

Proof: প্রমাণটা একেবারেই Schur decomposition-এর প্রমাণটার মতই, খালি এখানে কোনো complex number-এর ব্যাপার নেই। যদি তুমি সেই প্রমাণটা বুঝে থাকো, তবে এটা একেবারেই সহজ লাগবে। এখানে ছবিটা হল Fig 2, যেটা Fig 1-এর আদলে তৈরী। যদি ওই প্রমাণটা না পড়ে থাকো, তাও অসুবিধা নেই, এখানে বুঝে বুঝে পড়লেই হবে।

Shall show that there is a real, orthogonal P such that $P'AP$ is diagonal.

Shall use induction on n, the order of A.

<u>Basis</u>: If $n = 1$, then A is already diagonal.

So we can take $P = 1$.

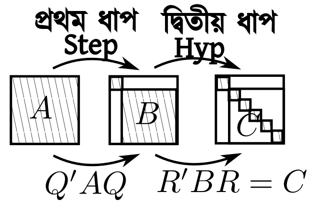


Fig 2

> Hypothesis: Assume the result for $n=1,...,m$.
>
> Step: Shall show for $n=m+1$.
>
> Let λ be an eigenvalue of A.
>
> $\because A$ is real, symmetric, $\therefore \lambda\in\mathbb{R}$.
>
> Then A has an eigenvector $\vec{v}\neq\vec{0}$ for λ with all entries real.

A একটা real matrix, আবার λ-ও একটা real সংখ্যা। কিন্তু তা বলে একটা eigenvector কেন থাকবেই, যার সব entry হবে real? তার কারণ এইরকম--

> ⟦ Because:
>
> Let $\vec{z}\in\mathbb{C}^n$ be any eigenvector of $A_{n\times n}$ for eigenvalue λ.

λ যখন eigenvalue সুতরাং eigenvector একটা থাকতে বাধ্য। তার entry-গুলো complex সংখ্যা হতে পারে। $\vec{z}$ এরকমই একটা eigenvector নিয়েছি।

> Then $A\vec{z}=\lambda\vec{z}$.
> $\therefore \mathrm{Re}(A\vec{z})=\mathrm{Re}(\lambda\vec{z})$ and $\mathrm{Im}(A\vec{z})=\mathrm{Im}(\lambda\vec{z})$.
> $\because A$ is real symmetric, $\therefore \lambda\in\mathbb{R}$.
> $\therefore A\,\mathrm{Re}(\vec{z})=\lambda\mathrm{Re}(\vec{z})$ and $A\,\mathrm{Im}(\vec{z})=\lambda\mathrm{Im}(\vec{z})$.

এখানে $\mathrm{Re}(\vec{z})$ আর $\mathrm{Im}(\vec{z})$ দুজনেরই সব entry হল real. মনে রেখো যে, $w=a+ib$ $(a,b\in\mathbb{R})$ হলে $\mathrm{Im}(w)=b$ (মোটেই ib নয়!)। এটাও দেখা গেল যে, $\mathrm{Re}(\vec{z})$ এবং $\mathrm{Im}(\vec{z})$ দুজনেই eigenvector হবার মূল শর্তটাও পালন করে। খালি nonzero কিনা সেটাই দেখা বাকি।

> Both $Re(\vec{z})$ and $Im(\vec{z})$ have all entries real.
> Since $\vec{z}\neq\vec{0}$, so at least one of $Re(\vec{z})$ and $Im(\vec{z})$ is nonnull, and hence is an eigenvector of A for eigenvalue λ.
>
> ⟧
>
> Let $\vec{u}=\frac{\vec{v}}{\|\vec{v}\|}$.
>
> We extend $\vec{u}$ to an ONB of $\mathbb{R}^n$ and let $Q_{n\times n}$ be the orthogonal matrix with these vectors as its columns.

মানে ONB-র vector-গুলোকে পাশাপাশি column হিসেবে দাঁড় করিয়ে Q বানিয়েছি। যেহেতু এর column-গুলো $\mathbb{R}^n$-এর ONB তৈরী করে, তাই Q একটা real, orthogonal matrix.

> Then $Q'AQ$ is symmetric with first column
>
> $$Q'A \times (\text{first column of } Q) = Q'A\vec{u} = Q'(\lambda\vec{u}) = \lambda Q'\vec{u} = \begin{bmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

এইটা ভালো করে বুঝে নাও।

> We partition $Q'AQ$ as
>
> $$Q'AQ = \begin{bmatrix} \lambda & \vec{0}' \\ \vec{0} & B \end{bmatrix},$$
>
> for some real, symmetric $(n-1)\times(n-1)$ matrix B.

B কেন real, symmetric হল? কারণ $Q'AQ$ নিজেই real, symmetric!

> By induction hypothesis, there is some real, orthogonal matrix R such that $R'BR = T$ is diagonal.
>
> Let $S = \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R \end{bmatrix}$.
>
> Then S is real, orthogonal.
>
> ⟦ <u>Because:</u>
>
> S is square and
>
> $$\begin{aligned} S'S &= \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R \end{bmatrix}' \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R \end{bmatrix} \\ &= \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & R'R \end{bmatrix} \\ &= \begin{bmatrix} 1 & \vec{0}' \\ \vec{0} & I_{n-1} \end{bmatrix} = I_n. \end{aligned}$$
>
> ⟧
>
> We define $P = QS$. Since product of orthogonal matrices is orthogonal, so P is orthogonal.
>
> Then $P'AP = S'(Q'AQ)S$ is diagonal, as required.
>
> So the result is proved by the principle of mathematical induction.

[Q.E.D]

Exercise 2: ধরো A একটা real, symmetric matrix, যার spectral decomposition হল $P'AP = D$. তবে দেখাও যে, D-এর diagonal entry-গুলো হল A-র eigenvalue, এবং P-এর column-গুলো হল A-র eigenvector-গুলো। ■

61.3.1 হাতে-কলমে

এবার ধরো তোমায় একটা real, symmetric matrix দিল, A. দিয়ে সত্যি সত্যি P আর D বার করতে বলল, যাতে $P'AP = D$ হয়, যেখানে P হল orthogonal এবং D হল diagonal. কী করে করবে? নীচের অংকটায় ঠিক সেটাই চেয়েছে।

Example 1: For the matrix $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$ find an orthogonal matrix P such that P^tAP is a diagonal matrix.[5] **(2021.CC12.3h)**

SOLUTION: আমাদের কাজ হল প্রথমে eigenvalue-গুলো বার করা, এবং তার পরে eigenvector-গুলো বার করা। যদি $n = 3$-খানা orthonormal eigenvector[1] পাই, তবেই আমরা D বার করে ফেলতে পারি, diagonal বরাবর eigenvalue-গুলোকে সাজিয়ে, আর corresponding eigenvector-গুলোকে পরপর column হিসেবে নিয়ে P বানিয়ে। ধাপে ধাপে এগোই, প্রথমে eigenvalue বার করা--

> The characteristic polynomial is
>
> $$|\lambda I - A| = \cdots = \lambda^3 - 12\lambda^2 + 36\lambda - 32 = (\lambda - 8)(\lambda - 2)^2.$$
>
> So eigenvalues are $8, 2, 2$.

এই determinant-টা বার করা, এবং সেখান থেকে eigenvalue-গুলো বার করা বেশ ঘাম-ঝরানো কাজ। যাই হোক, এবার eigenvector বার করার পালা--

> For eigenvalue 8, we solve the system $(8I - A)\vec{x} = \vec{0}$, ie,
>
> $$\begin{aligned} (8-4)x - 2y - 2z &= 0 \\ -2x + (8-4)y - 2z &= 0 \\ -2x - 2y + (4-4)z &= 0. \end{aligned}$$
>
> Solving we get $x = y = z$, and x is arbitrary. Thus the eigenspace is $\text{span}\{(1,1,1)\}$.
>
> Hence a unit eigenvector is
>
> $$\tfrac{1}{\sqrt{3}}(1,1,1).$$

এখানে একটাই eigenvector নিয়েছি (কারণ alg mult=geom mult=1).
এবার 2 eigenvalue-র জন্য eigenvector বার করতে হবে। এখানে দুটো eigenvector চাই, কারণ এখানে alg mult=geom mult=2.

---

[1] অর্থাৎ, eigenvector-গুলো সবাই unit vector এবং পরস্পরের সঙ্গে orthogonal.

For eigenvalue 2, we solve the system $(2I-A)\vec{x}=\vec{0}$, ie,

$$\begin{aligned}(2-4)x-2y-2z &= 0\\ -2x+(2-4)y-2z &= 0\\ -2x-2y+(2-4)z &= 0.\end{aligned}$$

Solving we get $z=-x-y$, where x,y are arbitrary. So the eigenspace is

$$\begin{aligned}&\{(x,y,-x-y)\ :\ x,y\in\mathbb{R}\}\\ =\ &\{x(1,0,-1)+y(0,1,-1)\ :\ x,y\in\mathbb{R}\}\\ =\ &\text{span}\{(1,0,-1),(0,1,-1)\}.\end{aligned}$$

Let

$$\begin{aligned}\vec{u}_1 &= (1,0,-1),\\ \vec{u}_2 &= (0,1,-1).\end{aligned}$$

Then $\{\vec{u}_1,\vec{u}_2\}$ is a basis of the eigenspace.

কারণ এটা একটা spanning set তো আছেই। এদিকে eigenspace-টার dimension হল 2 (যেহেতু $gm(2,A)=am(2,A)=2$)। আমাদের set-টা সাইজও সেটাই।
এবার এই basis-টাকে Gram-Schmidt লাগিয়ে ONB-তে পরিণত করব। প্রথমে orthogonal basis বানাই--

Applying Gram-Schmidt process to this basis we get an orthogonal basis as follows:

$$\begin{aligned}\vec{v}_1 &= \vec{u}_1=(1,0,-1)\\ \vec{v}_2 &= \vec{u}_2-\frac{\langle\vec{u}_2,\vec{v}_1\rangle}{\langle\vec{v}_1,\vec{v}_1\rangle}\vec{v}_1=\cdots=\tfrac{1}{2}(-1,2,-1).\end{aligned}$$

এবার orthogonal basis থেকে ONB পেতে আর কতক্ষণ?

Normalising,

$$\begin{aligned}\vec{w}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|}=\tfrac{1}{\sqrt{2}}(1,0,-1),\\ \vec{w}_2 &= \frac{\vec{v}_2}{\|\vec{v}_2\|}=\tfrac{1}{\sqrt{6}}(-1,2,1).\end{aligned}$$

We construct an orthogonal matrix P with these as its columns:

$$P=\left[\begin{array}{ccc}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}}\end{array}\right]$$

Then $P'AP = D$, where

$$D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

■

একইরকম আরেকটা অংক তোমার নিজের করার জন্য।

Exercise 3: Diagonalise the matrix $A = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{bmatrix}$ orthogonally.[5] **(2021.CC12.3a)**

■

DAY 62
Spectral decomposition-এর প্রয়োগ (part 1)

62.1 Quadratic form-এর maximum/minimum বার করা

আমরা ষষ্ঠ অধ্যায়ে লাট্টু ঘোরানোর উদাহরণটা দিয়েছিলাম, মনে আছে? কিংবা statistics-এর উদাহরণটা? দুই জায়গাতেই আমরা একটা quadratic form নিয়ে কাজ করছিলাম

$$\vec{w}'A\vec{w},$$

যেখানে $\vec{w}$ ছিল একটা unit vector, প্রশ্ন ছিল--এরকম কোন্ $\vec{w}$-এর জন্য quadratic form-টা maximum হবে? এর উত্তরটা আমরা তখনই উল্লেখ করেছিলাম--যখন $\vec{w}$ হবে A-র সবচেয়ে বড় eigenvalue-র জন্য একটা eigenvector. এটা তখন প্রমাণ করি নি, এবার করব। কিন্তু প্রথমে একটা অংক দেব, যেটা তোমাকে সহজবুদ্ধি ব্যবহার করে করতে হবে, eigenvalue, eigenvector ইত্যাদি ছাড়াই।

Exercise 4: যদি $M = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}$ হয় (ফাঁকা জায়গাগুলোয় শূন্য আছে), তবে $\vec{x}'M\vec{x}$ সবচেয়ে বেশী কত হতে পারে, যেখানে $\|\vec{x}\| = 1$?
HINT: এই অংকটা যেন বাদ দিয়ে যেও না। না পারলে উত্তর দেখে নিও। এই অংকের যুক্তিটা এক্ষুণি আমাদের কাজে লাগবে। ■

এবার তবে প্রমাণে হাত দেওয়া যাক। প্রমাণটা spectral decomposition-এর খুব গুরুত্বপূর্ণ একটা প্রয়োগ। কী প্রমাণ করতে চলেছি, সেটা আগে লিখে নিই--

THEOREM

Let A be a real, symmetric matrix. Then

1. the maximum possible value of the quadratic form $\vec{w}'A\vec{w}$ for $\vec{w} \in \mathbb{R}^n$ with $\|\vec{w}\| = 1$ is the largest eigenvalue of A.
2. this maximum value is attained if and only if $\vec{w}$ is a unit eigenvector of A corresponding to this eigenvalue.

এখানে unit eigenvector মানে এমন eigenvector যার norm হল 1.
Proof:

> (1) Let a spectral decomposition of A be
> $$A = PDP',$$
> where P is orthogonal and $D = diag(\lambda_1, ..., \lambda_n)$ with
> $$\lambda_1 \geq \cdots \geq \lambda_n.$$

এখানে আমরা "a spectral decomposition"লিখেছি, "the" লিখিনি, কারণ spectral decomposition মোটেই unique নয়। Eigenvalue-গুলোকে তুমি কাকে আগে কাকে পরে রাখবে, তার উপরে নির্ভর করে নানারকম D হতে পারে। আমরা এমনভাবে D নিলাম যাতে diagonal entry-গুলো বড় থেকে ছোটো সাজানো থাকে।

> Let $\vec{w} \in \mathbb{R}^n$ be any vector with $\|\vec{w}\| = 1$. Then
> $$\vec{w}'A\vec{w} = \vec{w}'PDP'\vec{w} = (\underbrace{P'\vec{w}}_{\vec{x}})'D(\underbrace{P'\vec{w}}_{\vec{x}}) = \vec{x}'D\vec{x} = \sum_1^n \lambda_i x_i^2,$$
> where $\vec{x} = (x_1, ..., x_n)'$.

আমরা $\|\vec{w}\| = 1$ নিয়ে শুরু করেছিলাম। Orthogonal matrix দিয়ে গুণ করলে norm যে বদলায় না, সেটাও আমরা জানি। সুতরাং $\vec{x} = P'\vec{w}$-এর norm-ও 1 হবে--

> Now
> $$\|\vec{x}\|^2 = \vec{x}'\vec{x} = \vec{w}'PP'\vec{w} = \vec{w}'\vec{w} = \|\vec{w}\|^2 = 1,$$
> since $PP' = I$.
> So $\sum x_i^2 = 1$.

তার মানে আমরা এমনভাবে x_i-দের নিতে চাইছি, যাতে $\sum x_i^2 = 1$ হয়, এবং সেই শর্তসাপেক্ষে $\sum \lambda_i x_i^2$ যথাসম্ভব বড় হয়।
এর জন্য আমরা যা করতে চলেছি, সেটা একটা গল্প দিয়ে বুঝে নেওয়া সহজ। গল্পটা আগের অংকের উত্তরেই একবার বলেছি। ধরো তুমি বাজারে চাল কিনতে গেছ। তোমার কাছে মোট যত টাকা আছে, সেটাকে 100% বা 1 ধরছি। মোট n রকমের চাল আছে। এর মধ্যে i-th চাল তুমি কিনবে x_i^2 টাকার। যেহেতু তুমি তোমার পুরো টাকাটাই খরচা করতে চাও, তাই $\sum x_i^2 = 1$.

এদিকে চালগুলোর দাম সবার একই নয়, i-th চালটা এক টাকায় পাওয়া যায় λ_i কেজি। সুতরাং তোমার কেনা মোট চালের পরিমাণ হবে $\sum \lambda_i x_i^2$. তুমি ঠিক করতে চাইছ, কোন্ চালের পিছনে কত টাকা ঢাললে এই পরিমাণটা সবচেয়ে বেশী হবে। এইভাবে ভাবলে উত্তরটা কি নেহাতই স্পষ্ট নয়? যে চালটা সবচেয়ে সস্তা, পুরো টাকাটা তো সেটাতেই ঢালা উচিত! এখানে যেহেতু λ_1 সবার বড়, মানে প্রথম চালটাই সবচেয়ে লাভজনক, তাই আমাদের উচিত $x_1^2 = 1$ এবং $x_2^2 = \cdots = x_n^2 = 0$ নেওয়া। সেক্ষেত্রে মোট লাভ হবে λ_1.
এই যুক্তিটাই এবার অংকের ভাষায় সাজিয়ে লিখব।

> Thus
> $$\sum_1^n \lambda_i x_i^2 \leq \sum_1^n \lambda_1 x_i^2 = \lambda_1. \qquad (*)$$

এখানে $\leq$-টা এল কারণ λ_1 হল λ_i-দের মধ্যে সবচেয়ে বড়, এবং x_i^2-রা সবাই ≥ 0.
সুতরাং quadratic form-টা λ_1-এর চেয়ে বেশী বড় হতে পারে না, এটুকু বোঝা গেল। এবার যদি দেখাতে পারি যে λ_1-এর সমান হতে পারে, তবেই λ_1-কে maximum-এর মর্যাদা দেওয়া যাবে।

> Also if $\vec{w}$ is an eigenvector corresponding to λ_1 then
> $$\vec{w}'A\vec{w} = \vec{w}'(\lambda_1 \vec{w}) = \lambda_1(\vec{w}'\vec{w}) = \lambda_1 \|\vec{w}\|^2 = \lambda_1.$$
> Thus the maximum possible value of the quadratic form is the largest eigenvalue λ_1, as required.

এবার দ্বিতীয় অংশ। এর একটা দিক তো প্রমাণ হয়েই গেল, $\vec{w}$ যদি λ_1-এর জন্য একটা unit eigenvector হয়, তবে quadratic form-টা maximum হবে।

> (2) The above arguement also proves that the maximum is attained when $\vec{w}$ is unit eigenvector corresponding to λ_1.

দেখানো বাকি আছে খালি উল্টো দিকটা, মানে quadratic form-টাকে maximum হতে গেলে $\vec{w}$-কে অবশ্যই λ_1-এর জন্য একটা unit eigenvector হতেই হবে।

> Conversely, let $\vec{w}$ be any unit vector for which the maximum is attained. Then equality holds in (*), and so
> $$\sum_1^n \lambda_i x_i^2 = \sum_1^n \lambda_1 x_i^2.$$
> Let $k \geq 1$ be such that
> $$\lambda_1 = \cdots = \lambda_k > \lambda_{k+1} \geq \cdots \geq \lambda_n.$$

যেমন λ_i-রা যদি $3, 3, 2, 2, 2, 1$ হয় তবে $\lambda_1 = \lambda_2 = 3 > 2 = \lambda_3$. সুতরাং $k = 2$.

> Then $x_{k+1} = \cdots = x_n = 0$.

এইটা কেন হল? ভেবে দেখো (*) কেন হয়েছিল--যেহেতু $\lambda_i \leq \lambda_1$. সুতরাং (*)-এ equality হবার জন্য যেসব $\lambda_i < \lambda_1$, তাদের কোনো প্রভাব থাকলে চলবে না, মানে ওদের বেলায় x_i-রা শূন্য হতে হবে।
এবার এ থেকে আমরা $\vec{w}$-এর বিষয়ে একটা সিদ্ধান্ত করব।

> We had $\vec{x} = P'\vec{w}$, and so $\vec{w} = P\vec{x}$, since P is orthogonal.
>
> Thus
> $$\vec{w} = x_1\vec{p}_1 + \cdots + x_n\vec{p}_n = x_1\vec{p}_1 + \cdots + x_k\vec{p}_k,$$
> since $x_{k+1} = \cdots = x_n = 0$.

এখানে সেই পূর্বপরিচিত জিনিসটা ব্যবহার করলাম, $P\vec{x}$ হল P-এর column-দের একটা linear combination, যেখানে coefficient-গুলো আসছে $\vec{x}$-এর entry-গুলো থেকে।

> Hence $A\vec{w} = \lambda_1(x_1\vec{p}_1 + \cdots + x_k\vec{p}_k) = \lambda_1\vec{w}$.

আর $\vec{w}$ তো unit vector নিয়েই শুরু করেছিলাম, তাই--

> Thus $\vec{w}$ must be a unit eigenvector of A corresponding to eigenvalue λ_1, as required.

[Q.E.D]

বাপ রে, কি লম্বা প্রমাণ! এই theorem-টা linear algebra-র বহুলব্যবহৃত প্রয়োগগুলোর অন্যতম।
একইরকম জিনিস আমরা minimum-এর জন্যও করতে পারি। সেটাই নীচের theorem-টার বক্তব্য।

THEOREM

Let A be a real, symmetric matrix. Then

1. the minimum possible value of the quadratic form $\vec{w}'A\vec{w}$ for $\vec{w} \in \mathbb{R}^n$ with $\|\vec{w}\| = 1$ is the smallest eigenvalue of A.
2. this minimum value is attained if and only if $\vec{w}$ is a unit eigenvector of A corresponding to this eigenvalue.

Exercise 5: আগের theorem-টার প্রমাণের কায়দায় এটাকেও প্রমাণ করো। তাতে আর কিছু না হোক, আগের theorem-টার প্রমাণটা ভালো করে শেখা হবে। ■

62.2 Definiteness-এর সঙ্গে সম্পর্ক

ষষ্ঠ অধ্যায়ে আমরা positive definite (pd), nonnegative definite (nnd) ইত্যাদি জিনিস নিয়ে আলোচনা করেছিলাম। এগুলো ছিল real symmetric matrix-দের শ্রেণীবিভাগ। একটা real symmetric matrix A-কে pd বলি যদি

$$\forall \vec{x} \in \mathbb{R}^n \quad \left[\vec{x} \neq \vec{0} \implies \vec{x}'A\vec{x} > 0\right].$$

একইভাবে $>$-এর জায়গায় $\geq$ বসালে হবে n.n.d., বা $<$ বসালে হবে negative definite (nd), ইত্যাদি। এই শ্রেণীবিভাগের সঙ্গে eigenvalue-দের একটা সহজ সম্পর্ক আছে--

THEOREM

A real, symmetric matrix is

- pd iff all its eigenvalues are > 0
- nnd iff all its eigenvalues are ≥ 0
- psd iff all its eigenvalues are ≥ 0, and at least one eigenvalue is 0.
- id iff at least one eigenvalue is > 0, and at least one eigenvalue is < 0.

Similar statements hold for nd, npd and nsd matrices with the inequalities reversed.

Proof: সবগুলোরই প্রমাণ একইরকম। আমরা খালি pd-র জন্য করি। প্রমাণটার দুটো দিক--এক, pd হলে eigenvalue-রা সকলেই > 0 হবে, আর দুই, eigenvalue-রা সবাই > 0 হলে matrix-টা pd হবে।

> (Only if part): Let A be a pd matrix. Let λ be any eigenvalue of A. Let $\vec{v} \neq \vec{0}$ be an eigenvector for this eigenvalue.
>
> $\because A$ is pd,
>
> $\therefore \vec{v}' A \vec{v} > 0,$
>
> $\therefore \vec{v}'(\lambda \vec{v}) > 0,$
>
> $\therefore \lambda \|\vec{v}\|^2 > 0,$
>
> $\therefore \lambda > 0,$ since $\|\vec{v}\|^2 > 0.$

এবার উল্টো দিকটা--

> (If part): Let all the eigenvalues of A be > 0.
>
> Take any $\vec{w} \in \mathbb{R}^n$ with $\vec{w} \neq \vec{0}$.

আমরা দেখাব যে, $\vec{w}' A \vec{w} > 0$ হবে।

> Let $A = PDP'$ be a spectral decomposition.
>
> Then
>
> $$\begin{aligned} \vec{w}' A \vec{w} &= \vec{w}' P D P' \vec{w} \\ &= \vec{x}' D \vec{x} \quad [\text{ where } \vec{x} = P' \vec{w}] \\ &= \sum \lambda_i x_i^2, \end{aligned}$$
>
> where $\vec{x} = (x_1, ..., x_n)'$ and $D = diag(\lambda_1, ..., \lambda_n)$. Here λ_i's are the eigenvalues of A.

সুতরাং $\vec{w}' A \vec{w} \geq 0$ হতে বাধ্য, কারণ λ_i-রা সকলেই > 0. কিন্তু আমাদের দেখানোর কথা ছিল $\vec{w}' A \vec{w} > 0$. তার জন্য দেখাতে হবে যে, অন্ততঃ একটা $x_i \neq 0$, মানে $\vec{x} \neq \vec{0}$.

Now $\vec{x} \neq \vec{0}$.

⟦ Because:

$\vec{w} \neq \vec{0}$, and P' is orthogonal ($\therefore$ nonsingular).
So $\vec{x} = P'\vec{w} \neq 0$.

⟧

So $\exists i \quad x_i \neq 0$. Also $\forall i \quad \lambda_i > 0$.
So $\sum \lambda_i x_i^2 > 0$.
Hence $\vec{w}' A \vec{w} > 0$, as required.

[Q.E.D]

Exercise 6: $A_{n\times n}$ is a negative definite matrix with $|A| = 3$. Then which of the following statements must be true?

(A) n is even (B) n is odd (C) n is divisible by 3 (D) n is not divisible by 3.

■

Exercise 7: এই matrix-টা কোন্ ধরণের definite?

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 5 \\ 0 & 0 & 6 \end{bmatrix}.$$

■

DAY 63
Spectral decomposition-এর প্রয়োগ (part 2)

63.1 Conic section-দের canonical form বার করা

Fig 3-এ কিছু conic section-এর নমুনা রয়েছে, circle, ellipse, parabola এবং hyperbola. এরা বিভিন্নভাবে ছড়িয়ে ছিটিয়ে রয়েছে। গ্রাফ দিয়ে এদের জ্যামিতিক বৈশিষ্ট্য আলোচনা করার জন্য সুবিধা হবে যদি আমরা axis-গুলোকে এমনভাবে নিই, যাতে conic section-গুলো দেখতে Fig 4, Fig 5 আর Fig 6-র মত সাজানো গোছানো হয়। এই "সাজানো গোছানো" রূপগুলোকে বলে **canonical form**. এর সঙ্গে ষষ্ঠ অধ্যায়ে শেখা quadratic form-এর canonical form-এর কোনো সম্পর্ক নেই। অবশ্য এখানেও equivalence-এর গল্প রয়েছে। দুটো conic section-কে আমরা equivalent ভাবছি, যদি ওদের একই আকৃতি হয় (অর্থাৎ কাগজ কেটে একটাকে আরেকটার উপর পুরো মিলিয়ে বসিয়ে দেওয়া যায়)। এই canonical form-এর সংজ্ঞাটা ভালো করে বুঝে নাও--

- Circle-এর বেলায় আমরা origin-টা নেব কেন্দ্রে। Fig 4 দেখো। Equation-টাকে লিখব $x^2 + y^2 - r^2 = 0$ আকারে।

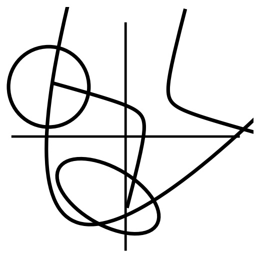


Fig 3

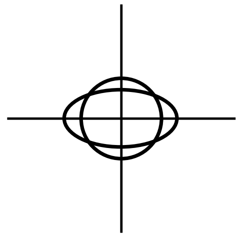


Fig 4

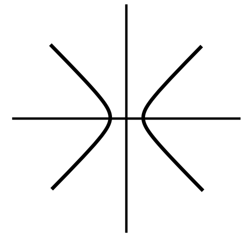


Fig 5

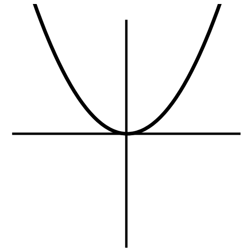


Fig 6

- Ellipse-এর বেলায় আমরা x-axis-টাকে নেব major axis বরাবর, আর origin-টা নেব কেন্দ্রে। Equation-টাকে লিখব--
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$
যেখানে $a > b > 0$. এই শেষের "$a > b$" শর্তটা ভুলে যেও না। এই জন্যই x-axis-টা থাকছে major axis বরাবর। Fig 4 দেখো।

- Hyperbola-র বেলাতেও x-axis-টাকে নেব major axis বরাবর, মানে যে axis-এর উপর focus দুটো আছে। আর origin যথারীতি কেন্দ্রে। Equation-টা নেব--
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$
এখানে কিন্তু a আর b-এর মধ্যে কে বড় কে ছোটো, তা নিয়ে আমাদের মাথাব্যথা নেই। Fig 5 দেখো।

- Parabola-র বেলায় আমরা origin-টাকে নেব vertex-এ। আর y-axis-টা নেব এমনভাবে যাতে খোলা দিকটা থাকে positive y-axis-এর দিকে। Equation-টাকে লিখব $y = ax^2$, যেখানে $a > 0$. বোঝার জন্য Fig 6 দেখো।

ছবি দিয়ে ভাবলে কোনো conic section-কে canonical form-এ এনে ফেলার জন্য axis দুটোকে খানিকটা ঘোরালে (অংকের ভাষায় rotation) আর খানিকটা সরালেই (মানে অংকের ভাষায় translation করলেই) হবে। Fig 7-এ ব্যাপারটা একটা ellipse-এর উদাহরণ দিয়ে বুঝিয়েছি। প্রথমে ellipse-টা ছিল গ্রাফ কাগজের axis দুটোর সাপেক্ষে হেলানো অবস্থায়, যেটাকে ছবিতে (A) বলেছি। (B) ছবিতে গ্রাফ কাগজের axis দুটোকে ঘুরিয়ে ellipse-এর axis দুটোর সঙ্গে সমান্তরাল করেছি। লক্ষ্য করো, এতে origin-এর স্থানপরিবর্তন হয়নি। শেষ ধাপে origin-টাকে সরিয়ে পৌঁছেছি (C)-এ, যেটা canonical form-এ আছে। এই ঘোরানো আর সরানোর ফলে কী হবে, সেটাই আমরা অংক কষে বার করতে চাই। মনে রেখো কতটা ঘোরানো বা কতটা সরানো হচ্ছে, সেটা নিয়ে আমাদের আগ্রহ নেই। শেষ পর্যন্ত conic section-টা কী রকম দেখাবে, সেটা নিয়েই আমাদের আগ্রহ।
ধরো তোমাকে একটা conic section-এর general equation দিলাম--
$$ax^2 + by^2 + cxy + px + qy + r = 0. \qquad (*)$$
এখানে x, y হল আমাদের variable দুটো, আর a, b, c, p, q, r হল বিভিন্ন সংখ্যা। এদের উপর নির্ভর করে এই equation-টা বিভিন্ন conic section দিতে পারে। অবশ্য conic section ছাড়া কিছু ভুলভাল জিনিসও ঢুকে পড়ে, যেমন ধরো $a = b = r = 1$ আর $c = p = q = 0$ নিলে equation-টা দাঁড়ায় $x^2 + y^2 + 1 = 0$, যেটা অসম্ভব (কারণ আমরা $x, y \in \mathbb{R}$ নিয়ে কাজ করছি)।
আমরা চাই কেবলমাত্র অংক কষে (কোনো ছবি ছাড়াই) (*)-কে canonical form-এ এনে ফেলতে। সেটা করতে গিয়ে conic section-টা কোন্ ধরণের সেটাও জানা হয়ে যাবে। ভাবতে পারো যেন (*) হল Fig 7-এর উদাহরণে (A) অবস্থানের গাণিতিক রূপ।

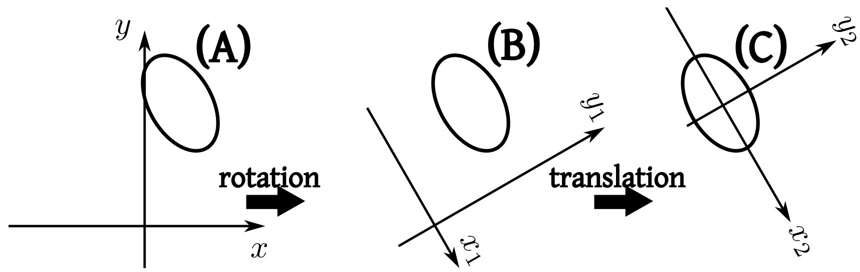


Fig 7

আমরা (*)-এর বাঁদিকটাকে এইভাবে matrix দিয়ে লিখে নেব--

$$\underbrace{ax^2+by^2+cxy}_{\text{quadratic}}+\underbrace{px+qy}_{\text{linear}}+r=\vec{x}'A\vec{x}+\vec{b}'\vec{x}+r,$$

যেখানে $\vec{x}=\begin{bmatrix} x \\ y \end{bmatrix}$, $A=\begin{bmatrix} a & \frac{c}{2} \\ \frac{c}{2} & b \end{bmatrix}$ এবং $\vec{b}=\begin{bmatrix} p \\ q \end{bmatrix}$. এটাকে আরো ছোটো করে লেখা যায়--

$$\begin{bmatrix} \vec{x}' & 1 \end{bmatrix}\underbrace{\begin{bmatrix} A & \vec{b} \\ \vec{b}' & r \end{bmatrix}}_{M}\begin{bmatrix} \vec{x} \\ 1 \end{bmatrix}.$$

আমাদের কাজে দুটো ধাপ-- অংক কষে axis দুটোকে ঘোরানো এবং তারপর সরানো। ঘোরানো মানে একটা orthogonal matrix দিয়ে গুণ করা। আর সরানো মানে কোনো vector যোগ করা। প্রথম ধাপে আমরা একটা orthogonal matrix নেব, P, এবং $\vec{x}=\begin{bmatrix} x \\ y \end{bmatrix}$-এর বদলে আমরা equation-টাকে $\vec{x}_1=\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ দিয়ে লিখব, যেখানে

$$\vec{x}=P\vec{x}_1.$$

আমাদের equation-এর quadratic অংশটা ছিল $\vec{x}'P\vec{x}$. এখন $\vec{x}$-এর বদলে $P\vec{x}_1$ লেখার ফলে সেটা দাঁড়াবে এইরকম--

$$\vec{x}_1'P'AP\vec{x}_1.$$

আমাদের উদ্দেশ্য হচ্ছে এমনভাবে P নেওয়া, যাতে x_1-axis আর y_1-axis হয় conic section-টার axis-এর সঙ্গে সমান্তরাল (Fig 7-এর (B) অবস্থানের মত)। তার জন্য x_1, y_1 দিয়ে লেখার পর equation-টায় কোনো x_1y_1 জাতীয় জিনিস থাকলে চলবে না, মানে ওই $P'AP$-কে diagonal হতে হবে। Spectral decomposition-এর গন্ধ পাচ্ছ? আমরা P নেব এমনভাবে, যাতে $P'AP=D$ হয়, যেখানে D হল diagonal. তার মানে $A=PDP'$ হবে A-র একটা spectral decomposition. অতএব D-এর diagonal entry-গুলো হবে A-র eigenvalue-গুলো, এবং P-এর column-গুলো হবে তাদের জন্য unit eigenvector.

যেই $P'AP$-টা diagonal হয়ে গেল, আর matrix ব্যবহার করার দরকার নেই। ধরো $D=\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$. তবে (*)-কে x_1, y_1 ব্যবহার করে লিখে নেওয়া যাবে এইভাবে--

$$\lambda x_1^2+\mu y_1^2+p_1x_1+q_1y_1+r=0, \qquad (**)$$

যেখানে p_1, q_1 কোনো দুটো উপযুক্ত সংখ্যা। যেহেতু x, y-এর বদলে x_1, y_1 দিয়ে লিখছি, তাই $px+qy$ পরিবর্তিত হয়ে $p_1x_1+q_1y_1$-এ পরিণত হয়েছে। আমাদের ঘোরানো পর্ব শেষ হল। একটা ব্যাপার মনে রেখো। এইভাবে লেখা সম্ভব বটে,

কিন্তু ওই p_1, q_1 বার করতে হলে তোমাকে P বার করতে হবে। কিন্তু পরের ধাপে p_1, q_1 কোথাও সরাসরি লাগবে না। তাই p_1, q_1 বার করব না, ফলে P-ও বার করতে হবে না। এই (**)-কে ভাবতে পারো Fig 7-এর (B) অবস্থানের equation.

এবার পরের ধাপ (Fig 7-এর (B) থেকে (C)-তে যাওয়া)। এই ধাপে আমরা গ্রাফ কাগজের axis দুটোকে এমনভাবে সরাতে চেষ্টা করব, যাতে ওই p_1x_1 আর q_1y_1 বিদায় হয়। এর জন্য আমরা square complete করতে চেষ্টা করব। আমাদের চেষ্টা থাকবে p_1x_1-কে λx_1^2-এর সঙ্গে, এবং q_1y_1-কে μy_1^2-এর সঙ্গে ভিড়িয়ে দিতে। এই কাজটা সবসময়েই করা যাবে যদি $\lambda, \mu \neq 0$ হয় (ওরা শূন্য হয়ে গেলে তো কোনো square-ই থাকছে না, তবে আর complete করবটা কাকে?)। Square complete করলে হবে

$$\lambda x_1^2 + p_1x_1 = \lambda \left(x_1 + \tfrac{p_1}{2\lambda}\right)^2 - \left(\tfrac{p_1}{2\lambda}\right)^2$$

আর

$$\mu y_1^2 + q_1y_1 = \mu \left(y_1 + \tfrac{q_1}{2\mu}\right)^2 - \left(\tfrac{q_1}{2\mu}\right)^2.$$

সুতরাং যদি $x_2 = x_1 + \frac{p_1}{2\lambda}$ আর $y_2 = y_1 + \frac{q_1}{2\mu}$ বসাই, তবে (**) হয়ে যাচ্ছে

$$\lambda x_2^2 + \mu y_2^2 + \underbrace{r - \left(\tfrac{p_1}{2\lambda}\right)^2 - \left(\tfrac{q_1}{2\mu}\right)^2}_{r_2}.$$

এটা হল Fig 7-এর উদাহরণে (C) অবস্থানের equation.

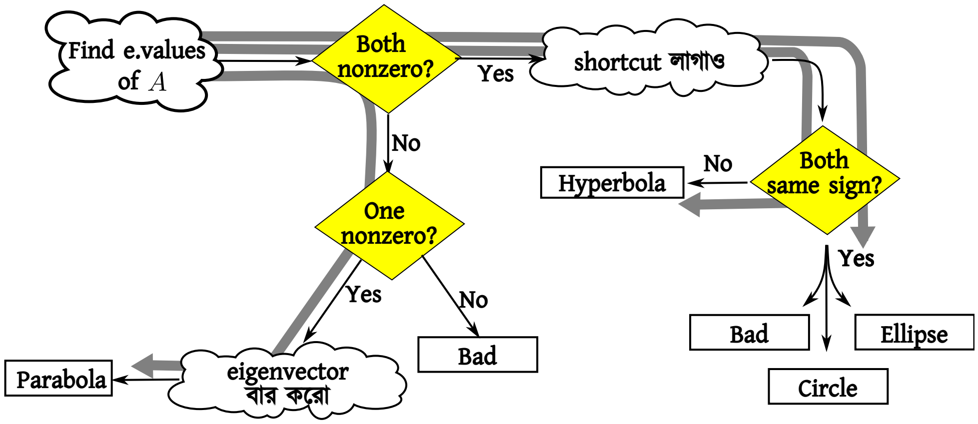
এর মধ্যে λ আর μ হল A-র eigenvalue দুটো, যেগুলো বার করা শক্ত নয়। খালি r_2-টাই এখনও অজানা। কিন্তু তার মধ্যে আবার p_1, q_1 রয়েছে, সেগুলো বার করতে গেলে P বার করতে হবে। এখানে একটা শর্টকাট আছে, যা দিয়ে p_1, q_1, P এসব কিছু মাত্র বার না করেই একবারে ধাঁ করে r_2 বার করে দেওয়া যাবে। শর্টকাটটা হল--

$$r_2 = \frac{|M|}{|A|}.$$

M কোন্ matrix-টা ছিল, মনে আছে তো? আমরা যখন (*) একটা 3×3 সাইজের matrix দিয়ে লিখেছিলাম, তখন এই matrix-টা পেয়েছিলাম। এই শর্টকাটটা কেন কাজ করে, সেটা একেবারে শেষে গিয়ে বলব।

এতক্ষণ যে আলোচনাটা করলাম, সেটা হল যখন λ আর μ দুজনেই nonzero. যদি λ আর μ-এর মধ্যে কোনো একজন শূন্য হয়ে যায়, তবে তার জন্য আর square complete-এর কৌশলটা কাজ করবে না। সেখানে একটু বাড়তি খাটতে হবে। সেই জটিলতায় একটু পরে আসব।

এই হচ্ছে মূল ব্যাপারটা। পুরো প্রক্রিয়াটার মধ্যে ছোটোখাটো বেশ কিছু খুঁটিনাটি আছে। তার কারণ হল (*)-এর মধ্যে conic section ছাড়াও কিছু ভুলভাল জিনিস ঢুকে গেছে, সেগুলোকে চিনে চিনে বাদ দিতে হবে, কারণ ওদের তো আর canonical form হয় না। সবগুলো ধাপ একসঙ্গে একটা flow chart-এর মত করে দিচ্ছি। তারপর উদাহরণ দিয়ে বোঝাব।



এখানে আমরা শেষ পর্যন্ত পাচ্ছি circle বা ellipse বা hyperbola বা parabola বা bad. ওই bad মানে হল সেই ভুলভাল জিনিসগুলো। Flowchart-এর গায় আমরা তিনটে পথ দেখিয়েছি মোটা করে। যদি ওইসব bad কেসগুলো না আসে, তবে তুমি ওই তিনটের মধ্যে একটা দিয়ে যেতে বাধ্য। সবচেয়ে ডানদিকের পথটা আমাদের নিয়ে ফেলে circle বা ellipse-এ, মাঝেরটা hyperbola-তে, আর বাঁদিকেরটা parabola-য়। আমরা একে একে এই তিন ধরণের উদাহরণ দেখব। প্রথম ধাপ সবক্ষেত্রেই A-র eigenvalue-গুলো বার করা।
শুরু করি সবচেয়ে ডানদিকের পথের উদাহরণ দিয়ে--

Example 2: Reduce the equation $7x^2 - 2xy + 7y^2 - 16x + 16y - 8 = 0$ into canonical form and determine the nature of the conic.[5] **(2020.CC12.3a)**

SOLUTION:

> The quadratic form in the given equation is
>
> $$\begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} 7 & -1 \\ -1 & 7 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \end{bmatrix}.$$

এবার A-র eigenvalue বার করব--

> The eigenvalues of A are 8 and 6.

এখানে ধাপগুলো আর দেখালাম না। ওগুলো তুমি ঢুকিয়ে নিও। যেহেতু eigenvalue দুটোই nonzero এবং ওদের একই চিহ্ন, এবং ওরা অসমান, তাই বোঝাই যাচ্ছে যে, আমরা ellipse পেতে চলেছি (যদি না ভুলভাল কিছু হয়)।

> Let a spectral decomposition of A be
>
> $$A = P \begin{bmatrix} 6 & 0 \\ 0 & 8 \end{bmatrix} P'.$$

যেহেতু ellipse, তাই দুটো eigenvalue-এর মধ্যে যেটা absolute value-তে ছোটো, সেটাকে উপরে রেখেছি। এর ফলে x_1-axis-টা থাকবে major axis-এর সঙ্গে সমান্তরাল।

> So after rotating (x, y) to (x_1, y_1), where $\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, we can write the equation as
> $$6x_1^2 + 8_1 y^2 + p_1 x_1 + q_1 y_1 - 8 = 0$$
> for some $p_1, q_1 \in \mathbb{R}$.

এবার আমাদের উচিত গ্রাফ কাগজের axis দুটোকে সরিয়ে origin-টাকে ellipse-টার কেন্দ্রে আনা।

> Translating the coordinate system from (x_1, y_1) to (x_2, y_2), we can write this as $6x_2^2 + 8y_2^2 + r_2 = 0$,

Flowchart বলেছে এবার সেই শর্টকাটটা লাগাতে হবে--

> where
> $$r_2 = \frac{\begin{vmatrix} 7 & -1 & -8 \\ -1 & 7 & 8 \\ -8 & 8 & -8 \end{vmatrix}}{|A|} = \frac{-1152}{48} = -24.$$

এবার সাজিয়ে লিখলেই ellipse বুঝতে পারবে--

> So we have $6x_2^2 + 8y_2^2 = 24$, or
> $$\frac{x_2^2}{4} + \frac{y_2^2}{3} = 1,$$
> which is the canonical form of an ellipse with semimajor axis 2 and semiminor axis $\sqrt{3}$.

■

এই অংকটায় আমরা দুই জায়গায় পরিশ্রম বাঁচিয়েছি--

- এক, সমাধানে P-এর উল্লেখ করলেও সেই matrix-টা কিন্তু সত্যি সত্যি বার করি নি।
- আর দুই, যদি Fig 7-এর সঙ্গে তুলনা করে ভাবো, তবে আমরা (B) অবস্থানের equation-টা পুরো বার করিনি। শর্টকাট ব্যবহার করে (A) অবস্থান থেকে সোজা (C) অবস্থানের equation-টায় গিয়ে পড়েছি।

অবশ্য কতটা ঘোরানো আর কতটা সরানো হয়েছে, সেটাও যদি প্রশ্নে জানতে চাইত, তবে আমাদেরকে পুরো P-ই বার করতে হত, এবং সেটা অনেক বেশী পরিশ্রমের কাজ হত।
আমরা flowchart-এর সবচেয়ে ডানদিকের পথের একটা উদাহরণ দেখলাম। এবার মাঝের পথের একটা অংক দিই। কায়দাটা একেবারেই একইরকম, তাই আর কষে দেব না।

Exercise 8: Find the canonical form of $-x^2 - y^2 - 14xy - 44x - 20y - 64 = 0$. Also find the nature of the conic section.
■

এবার flowchart-এর সবচেয়ে বাঁদিকের পথের একটা উদাহরণ দেখি, যেখানে একটা eigenvalue হবে zero আর অন্যটা nonzero.

Example 3: Reduce the equation $9x^2 - 24xy + 16y^2 + 2x - 11y + 16 = 0$ to its canonical form and determine the nature of the conic.[5] **(2022.CC12.3h)**

SOLUTION:

> Here the quadratic part is $\vec{x}'A\vec{x}$, where
> $$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} \text{ and } A = \begin{bmatrix} 9 & -12 \\ -12 & 16 \end{bmatrix}.$$

শুরু করব যথারীতি A-র eigenvalue বার করা দিয়ে--

> A has eigenvalues 25 and 0.

একজন zero, অন্যজন nonzero. তাই এখানে আমরা সবচেয়ে বাঁদিকে পথের পথিক। অবশ্য এর পরের কাজটা সব পথেই একই, x_1, y_1 দিয়ে লেখা, যেখানে কোনো x_1y_1 থাকবে না।

> Let a spectral decomposition of A be
> $$A = P \begin{bmatrix} 6 & 0 \\ 0 & 8 \end{bmatrix} P'.$$
> So after rotating (x, y) to (x_1, y_1), where $\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, we can write the equation as
> $$25x_1^2 + p_1x_1 + q_1y_1 + 16 = 0,$$
> for some $p_1, q_1 \in \mathbb{R}$.

আমরা square complete করে p_1x_1-কে $25x_1^2$-এর সঙ্গে মিশিয়ে লোপাট করে দিতে পারব। এইটুকু আগের অংকের মতই। কিন্তু যেহেতু কোনো y_1^2 নেই, তাই q_1y_1-কে লোপাট করা যাবে না। তাই q_1 বার করতেই হবে। তার জন্য y_1-কে x আর y দিয়ে লিখতে হবে। এখানে $\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ ছিল। তার মানে, P-টা orthogonal হওয়ায়

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = P' \begin{bmatrix} x \\ y \end{bmatrix}.$$

সুতরাং $y_1 = \alpha x + \beta y$, যেখানে $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ হল P-র দ্বিতীয় column, মানে 0-র জন্য একটা unit eigenvector. এই জায়গাটায় মাথা গুলিয়ে যাওয়া স্বাভাবিক। ভালো করে কয়েকবার পড়ে বুঝে নাও। তাই এইবার আমাদের একটা eigenvector-ও বার করতে হবে, 0-র জন্য। খানিকটা অংক কষলেই দেখবে এরকম একটা eigenvector হল $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$. একে unit vector-এ পরিণত করলে হয় $\frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix}$. এই eigenvector বার করার ধাপগুলো আর দেখালাম না, তুমি নিজেই ঢুকিয়ে নিও।

> One unit eigenvector corresponding to the eigenvalue 0 is $\begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$. So we can take
> $$y_1 = \frac{4}{5}x + \frac{3}{5}y.$$

শুরুতে y-ওয়ালা term-টা ছিল $-11y$. সেটা থেকেই এসেছে এই $q_1 y_1 = q_1\left(\frac{4}{5}x + \frac{3}{5}y\right)$. তার মানে $-11 = \frac{3q_1}{5}$, বা $q_1 = -\frac{55}{3}$.

> Now, matching coefficient of y in $-11y$ and $q_1 y_1 = q_1\left(\frac{4}{5}x + \frac{3}{5}y\right)$, we get $q_1 = -\frac{55}{3}$.

তার মানে আমাদের equation-টা এখন আছে

$$25x_1^2 + p_1 x_1 - \frac{55}{3}y_1 + 16 = 0.$$

এবার translation করার সময়ে $p_1 x_1$-কে $25x_1^2$-এর সাথে, এবং শেষে যা constant পড়ে থাকবে, সেটাকে $-\frac{55}{3}y_1$-এর সাথে মিশিয়ে লোপাট করে দেব।

> So after translation we can write the equation as $25x_2^2 - \frac{55}{3}y_2 = 0$, which has the canonical form
> $$y_2 = \frac{15}{11}x_2^2.$$

■

এই কটা অংক থেকে যা শিখলে, সেটা প্রয়োগ করে নীচের অংকটা করো।

Exercise 9: Find the canonical form and nature of each of the following conic sections:

1. $-y^2 - 14\,x\,y + 2\,y - x^2 + 14\,x - 13 = 0$
2. $19\,y^2 + 4\,x\,y + 76\,y + 16\,x^2 + 8\,x + 64 = 0$
3. $11\,y^2 - 6\,x\,y + 16\,y + 19\,x^2 + 32\,x + 19 = 0$
4. $-6\,y^2 - 12\,x\,y - 49\,y - 6\,x^2 - 47\,x - 98 = 0$
5. $7\,y^2 + 2\,x\,y + 44\,y + 7\,x^2 + 20\,x + 88 = 0$

■

63.1.1 শর্টকাটটা কেন কাজ করে?

এইবার বলি কেন আমাদের শর্টকাটটা কাজ করে। আমরা শুরু করেছিলাম এইরকম একটা general চেহারার equation থেকে--

$$ax^2 + by^2 + cxy + px + qy + r = 0.$$

এটাকে আমরা matrix দিয়ে লিখেছিলাম এইভাবে--

$$\begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} a & c/2 & p/2 \\ c/2 & b & q/2 \\ p/2 & q/2 & r \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}.$$

এই যে মাঝখানের 3×3 সাইজের matrix-টা, এটার একটা আশ্চর্য গুণ আছে। তুমি গ্রাফ কাগজটাকে ঘোরালে বা সরালে matrix-টা বদলে যেতেই পারে, কিন্তু ওর determinant-টা একই থাকতে বাধ্য। একটা উদাহরণ নিয়ে দেখাই।

Example 4: ধরো equation-টা আছে $x^2 + 2y^2 + 10 = 0$. তবে matrix-টা হবে

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

এর determinant হল 20. এবার ধরো আমরা x-এর বদলে $x + 2$ আর y-এর জায়গায় $y - 1$ বসালাম, তবে নতুন matrix-টা বার করো, এবং পরীক্ষা করে দেখো determinant-টা একই আছে কিনা।

SOLUTION: পরিবর্তিত equation-টা হবে $(x+2)^2 + 2(y-1)^2 + 10 = 0$, বা $x^2 + 2y^2 + 4x - 4y + 16 = 0$. সুতরাং নতুন matrix-টা হচ্ছে--

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & -2 \\ 2 & -2 & 16 \end{bmatrix}.$$

এর determinant-ও হল সেই 20. ■

এমনটা যে সবসময়েই হবে, সেটা সহজে প্রমাণ করার একটা কৌশল আছে। তুমি যদি গ্রাফ কাগজটাকে $\vec{a}$ পরিমাণ সরাও, মানে $\vec{x}$-এর জায়গায় $\vec{x} + \vec{a}$ বসাও, তবে সে কাজটাকে একটা matrix-এর গুণ দিয়ে প্রকাশ করা যায়--

$$\begin{bmatrix} \vec{x} + \vec{a} \\ 1 \end{bmatrix} = \begin{bmatrix} I_2 & \vec{a} \\ \vec{0}' & 1 \end{bmatrix} \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix}.$$

একইভাবে যদি P একটা 2×2 সাইজের matrix হয়, তবে সেই matrix দিয়ে $\vec{x}$-কে গুণ করাকে এইভাবে লেখা যায়--

$$\begin{bmatrix} P\vec{x} \\ 1 \end{bmatrix} = \begin{bmatrix} P & \vec{0} \\ \vec{0}' & 1 \end{bmatrix} \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix}.$$

এই ব্যাপারটা computer graphics-এ খুব ব্যবহার হয়। আমরা এখানে এইটা ব্যবহার করে দেখাব conic section-এর matrix-টার determinant কেন বদলায় না। ধরো শুরুতে matrix-টা ছিল M, মানে equation-টা ছিল--

$$\begin{bmatrix} \vec{x}' & 1 \end{bmatrix} M \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix} = 0.$$

এবার ধরো আমরা গ্রাফ কাগজটাকে $\vec{a}$ পরিমাণ translate করলাম, মানে $\vec{x}$-এর জায়গায় $\vec{x} + \vec{a}$ বসালাম। এর ফলে নতুন equation-টা হল

$$\begin{bmatrix} (\vec{x} + \vec{a})' & 1 \end{bmatrix} M \begin{bmatrix} (\vec{x} + \vec{a}) \\ 1 \end{bmatrix} = 0. \qquad (*)$$

এদিকে সেই matrix গুণের কৌশলটা লাগিয়ে আমরা লিখতে পারি--

$$\begin{bmatrix} \vec{x} + \vec{a} \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} I_2 & \vec{a} \\ \vec{0}' & 1 \end{bmatrix}}_{T} \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix}.$$

তাই (*)-টা দাঁড়াচ্ছে--

$$\begin{bmatrix} \vec{x}' & 1 \end{bmatrix} T'MT \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix} = 0.$$

সুতরাং নতুন matrix-টা হল $T'MT$. এবার লক্ষ্য করো যে, $|T| = 1$. তাই

$$|T'MT| = |T'| \cdot |M| \cdot |T| = |M|,$$

অর্থাৎ translation-এর বেলায় নতুন matrix-এর determinant-টা শুরুর matrix-টার determinant-এর সমান হচ্ছে।

Exercise 10: একইভাবে rotation-এর বেলাতেও দেখাও।
HINT: Rotation matrix-দের determinant কত হয়? ■

এইবার দেখি শর্টকাটটা কেন কাজ করছিল। আমরা গ্রাফ কাগজ ঘুরিয়ে সরিয়ে যে canonical form-এ পৌঁছেছিলাম, সেটার matrix-টা ছিল এইরকম--

$$\begin{bmatrix} D & \vec{0} \\ \vec{0}' & r_2 \end{bmatrix}.$$

শুরুর matrix-টার নাম দিয়েছিলাম M. এদের দুজনের determinant সমান। তাই

$$\begin{vmatrix} D & \vec{0} \\ \vec{0}' & r_2 \end{vmatrix} = |M|.$$

কিন্তু বাঁদিকের matrix-টা যেহেতু diagonal, তাই ওর determinant-টা হবে $r_2|D| = r_2|A|$, কারণ $|D|$ হল A-র eigenvalue দুটোর গুণফল, এবং সেটাই হল $|A|$. তাই

$$r_2|A| = |M|.$$

আমরা শর্টকাটটা ব্যবহার করছিলাম যখন দুটো eigenvalue-ই nonzero ছিল, তাই $|A| \neq 0$. সুতরাং পাচ্ছি--

$$r_2 = \frac{|M|}{|A|},$$

এবং সেটাই ছিল আমাদের শর্টকাট।

DAY 64
Singular value decomposition (SVD)

এবার একটা জিনিস শিখব, যেটা আজকের neural network আর machine learning-এর যুগে ভীষণই গুরুত্বপূর্ণ হয়ে উঠেছে। জিনিসটা হল spectral decomposition-এরই বড় ভাই, নাম **singular value decomposition (SVD)**. অনেক সিলেবাসেই এটা এখনো স্থান পায়নি। যারা কেবলমাত্র সেইসব সিলেবাসের পরীক্ষায় পাশের জন্য এ বই পড়ছ, তাদের জোর করব না, কিন্তু আধুনিক যুগে SVD না জেনে linear algebra পড়া, আর মোবাইল ফোন ব্যবহার না করে জীবনধারণ করা কতকটা একইরকম। তাই নেহাত আপত্তি না থাকলে আজকের পড়ায় অন্ততঃ একটু চোখ বুলিও।

64.1 কী জিনিস?

আমরা দেখেছি যে, যদি A একটা real, symmetric matrix হয়, তবে তার spectral decomposition হবে--

$$A = PDP',$$

যেখানে D হল diagonal এবং P হল orthogonal. চেহারাটা লক্ষ্য করো, A-কে তিনটে matrix-এর গুণফল হিসেবে লেখা হয়েছে, প্রান্তের দুটো matrix হল orthogonal, এবং মাঝেরটা diagonal. প্রান্তের matrix দুটো পরস্পরের transpose-ও বটে। SVD-ও অনেকটা একইরকম, এখানেও একটা matrix-কে তিনটে matrix-এর গুণফল হিসেবে লেখা হবে--

$$A = USV'.$$

এখানেও দুপ্রান্তের matrix দুটো হবে orthogonal এবং মাঝেরটা diagonal. কিন্তু spectral decomposition-এর সঙ্গে এর প্রধান পার্থক্য হল--

- এর জন্য A-কে symmetric হতে হবে না, যেকোনো real matrix হলেই[2] চলবে। এমন কি matrix-টার square হবারও কোনো প্রয়োজন নেই।
- দ্বিতীয়তঃ, এখানে দু প্রান্তের orthogonal matrix দুটো পরস্পরের transpose নাও হতে পারে। এমন কি ওদের সাইজও আলাদা হতে পারে।

গুছিয়ে লিখলে--

SVD (over $\mathbb{R}$)

Any real $A_{m\times n}$ can be factorised as

$$A = USV'$$

for some orthogonal $U_{m\times m}$ and $V_{n\times n}$ and diagonal $S_{m\times n}$ such that for some integer $r \geq 0$, the first r diagonal entries of S are positive real numbers in non-increasing order, the rest being zeroes.

সবটা পড়তে গিয়ে মাথা গুলিয়ে যাওয়া স্বাভাবিক। ছবি দিয়ে ভাবলে সুবিধা হবে। যদি A-টা square হয়, মানে $m = n$ হয়, তবে ব্যাপারটা Fig 8-এর মত। এখানে সাদা অংশগুলো শূন্য দিয়ে ভরা। S-এর diagonal বরাবর যে অংশটা শেড করা, সেখানকার সবগুলো entry-ই positive, এবং বড় থেকে ছোটো সাজানো। যদি $m > n$ হয়, তবে A-টা লম্বা খাড়া, সেক্ষেত্রে SVD-টা দেখতে Fig 9-র মত। আর যদি $m < n$ হয়, তবে দেখতে হবে Fig 10-র মত। আরো লক্ষ্য করো, S আর A-র সাইজ সবক্ষেত্রেই একই।

এই theorem-টা, এটা কেবলমাত্র SVD-র existence-এর কথা বলছে, uniqueness-এর কোনো দাবী করছে না। বস্তুতঃ SVD মোটেই unique নয়। সেটা খুব সহজেই দেখা যায়--যদি $A = USV'$ একটা SVD হয়, তবে U-এর জায়গায় $-U$ এবং V-এর জায়গায় $-V$ বসালে $A = (-U)S(-V)'$ হবে আরেকটা SVD.

Example 5: নীচে কিছু S দিলাম। তোমাকে বলতে হবে এদের মধ্যে কোনগুলো SVD-তে আসতে পারে। এবং সেক্ষেত্রে m, n আর r কত।

[2]বস্তুতঃ complex matrix-এর বেলাতেও SVD সম্ভব। সেটা আমরা পরে বলব।

Fig 8

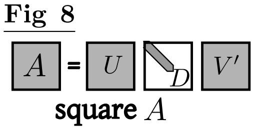


Fig 9

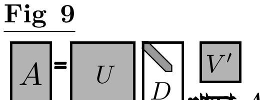
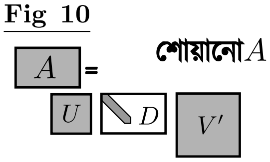


Fig 10



(1) $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ **(2)** $\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ **(3)** $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ **(4)** $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

SOLUTION:

1. এখানে আগে 1, পরে 2. কিন্তু আমাদের দরকার S-এর diagonal entry-গুলো যেন বড় থেকে ছোটো সাজানো থাকে। তাই এটা চলবে না।
2. এটাও চলবে না, কারণ diagonal-এ negative থাকা নিষেধ।
3. এটা চলবে। এখানে $m = n = r = 2$.
4. এটাও চলবে। এখানে $m = 4$, $n = 3$ এবং $r = 2$.

■

Exercise 11: নীচে কিছু S দিলাম। তোমাকে বলতে হবে এদের মধ্যে কোনগুলো SVD-তে আসতে পারে। এবং সেক্ষেত্রে m, n আর r কত।

(1) $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ **(2)** $\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$ **(3)** $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ **(4)** $\begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$ **(5)** $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

■

এই প্রসঙ্গে কয়েকটা ভাষা শিখে রাখো। ধরো S-এর i-th diagonal entry-টাকে σ_i বললাম, আর U এবং V-এর i-th column-গুলোকে যথাক্রমে $\vec{u}_i$ এবং $\vec{v}_i$ নাম দিলাম। তবে σ_i-কে বলব A-র i-th **singular value.** আর $\vec{u}_i$ হবে তার জন্য একটা **left singular vector** এবং $\vec{v}_i$ হবে তার জন্য একটা **right singular vector**.

> DEFINITION: **singular value, left/right singular vector**
>
> Let $A_{m\times n}$ be any matrix with an SVD $A = USV'$. Then the i-th diagonal entry of S is called the i-th **singular value** of A. The i-th column of U is called a corresponding **left singular vector** and the i-th column of V is called a corresponding **right singular vector** of A.

লক্ষ্য করো singular value-র আগে "the" ব্যবহার করেছি, কিন্তু left/right singular vector-দের বেলায় "a". এর কারণ হল তুমি যাই SVD নাও না কেন, singular value-গুলো একই থাকবে, কিন্তু বিভিন্ন SVD-র জন্য বিভিন্ন left/right singular vector আসতেই পারে।

পুরো ব্যাপারটা অনেকটা eigenvalue আর eigenvector-এর মত। খালি সেখানে eigenvalue-দের সঙ্গে আমরা eigen-vector-দের জোড় বাঁধতাম, আর এখানে প্রতিটা singular value-র সঙ্গে দুটো vector-এর জোড় বাঁধছি--একটা left singular vector আর একটা right singular vector. আরো একটা পার্থক্য আছে। eigenvector-রা nonzero vector হলেই চলত, কিন্তু singular vector-দের unit vector হতে হয়।

একটা উদাহরণ দেখা যাক।

Example 6: ধরো $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. এর জন্য কয়েকটা SVD এরকম--

1. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

2. $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

3. $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

এদের থেকে কী কী singular value আর left/right singular vector পাওয়া যাচ্ছে?

SOLUTION: সবক্ষেত্রেই মাঝের diagonal matrix-টা একই। তাই দুটো singular value-ই হল 1.

প্রথম ক্ষেত্রে left singular vector দুটো হল $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ এবং $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Right singular vector দুটোও তাই।

দ্বিতীয় ক্ষেত্রে একটা অন্য left/right singular vector এসেছে, $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

তৃতীয় ক্ষেত্রে left/right singular vector-রা হল $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ আর $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$. ■

এই উদাহরণে left/right singular vector-রা সমান হচ্ছিল, কারণ matrix-টা identity নিয়েছিলাম। অন্য উদাহরণে ওরা আলাদা হতেই পারে। যেমন নীচের অংকটায়।

Exercise 12: এই matrix-টার একটা SVD দাও, এবং তার ভিত্তিতে singular value-গুলো এবং তাদের জন্য left singular vector আর right singular vector বার করো--

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

■

Exercise 13: দেখাও যে, যেকোনো orthogonal matrix P নিলেই identity matrix-এর জন্য একটা SVD হয় $I = PIP'$. ■

Exercise 14: দেখাও যে, A-র singular value-রা হল $A'A$-র positive eigenvalue-দের positive square root. এ থেকেই বুঝবে কেন singular value-গুলো বিভিন্ন SVD-র জন্য আলাদা আলাদা হয় না।

HINT: যদি A-র একটা SVD নাও $A = USV'$, তবে $A'A$ কত হবে? ■

Exercise 15: যদি আগের অংকে $A'A$-র বদলে AA' নিতে, তবেও কি চলত? ■

এই অংক দুটো করার পর নীচের অংকটায় হাত দাও--

Exercise 16: $A_{1\times 3} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ এবং $B_{3\times 1} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}$-এর singular value-গুলো বার করো। ■

নীচের theorem-টা খুবই সহজ--

THEOREM

A matrix with rank k has exactly k singular values (not necessarily distinct).

Proof: যদি $A = USV'$ হয় একটা SVD, তবে $r(A) = r(S)$, কারণ U আর V হল nonsingular, তাই ওদের দিয়ে গুণ করলে rank বদলাবে না। একটা diagonal matrix-এর rank হয় তার nonzero diagonal entry-র সংখ্যা। এবং সেই সংখ্যাটাই আবার A-র singular value-দের সংখ্যাও বটে। **[Q.E.D]**

64.2 একটা সহজ প্রয়োগ

SVD-র একটা সহজ প্রয়োগ দেখা যাক। এটা খুব যে কাজের জিনিস এমন নয়, কিন্তু ছবি দিয়ে বোঝার পক্ষে মজাদার। যারা কম্পিউটারে বা মোবাইলে একটু-আধটু ছবি এডিট করতে জানো, তারা নিশ্চয়ই দেখেছ যে, একটা ছবিকে ছোটো বড় করা, বা খানিকটা ঘোরানো, বা ওল্টানো, এইসব কাজ খুব কঠিন কিছু নয়। মোবাইলের পর্দায় দুআঙুলে চিমটি কেটে ছবিকে ছোটো বড় করা তো আমরা আকছারই করে থাকি। কিন্তু অনেক সময়ে এমন কিছু করার দরকার পড়ে, যার জন্য কোনো সহজ পথ চোখে পড়ে না। একটা উদাহরণ দেখিয়েছি Fig 11-এ। ছবিটায় একটা বাক্স আর একটা মুখ দেখিয়েছি। ধরো এগুলো ইন্টারনেট থেকে ডাউনলোড করা দুটো ছবি। এবার আমরা চাই ওই মুখটাকে বাক্সের উপরে সেঁটে নীচের ছবিটা বানাতে। যেহেতু বাক্সের উপরের পাশটা আমরা তেরছাভাবে দেখছি , তাই সাঁটার সময়ে মুখটাকে কিন্তু কায়দা করে বাঁকিয়ে হেলিয়ে নিতে হয়েছে। এবং সেখানেই আমাদের প্রশ্ন, আমরা তো খালি ছোটো-বড় করা, ঘোরানো আর ওল্টানো পারি। সেগুলো কোনোভাবে মিলিয়েমিশিয়ে কি এই জটিল বাঁকানো আর হেলানো সম্ভব?

এখানে অপ্রত্যাশিতভাবে SVD আমাদের কাজে আসবে। আমরা জানি যে, ঘোরানো আর ওল্টানো মানে হল orthogonal matrix দিয়ে linear transformation করা। একইভাবে ছোটোবড় করা হল diagonal matrix দিয়ে গুণ, যেখানে diagonal entry-গুলো positive. এবার লক্ষ্য করো, আমরা যে জিনিসটা করতে চাইছি সেটাও একটা linear transformation, কারণ বাক্সের উপরটা সমতল, ফলে ওর গায়ে সাঁটার ফলে parallelogram-রা parallelgram-ই থাকবে। তার মানে আমাদের প্রশ্নটা দাঁড়াচ্ছে এই--

> আমাদের linear transformation-টাকে কিছু orthogonal matrix আর diagonal matrix মিলিয়ে বানানো যায় কিনা।

SVD বলছে যে, আলবৎ যায়, কারণ যেকোনো linear transformation-কেই একটা matrix দিয়ে লেখা যায়, এবং সেটাকে দুটো orthogonal matrix আর একটা diagonal matrix-এর গুণফল হিসেবে প্রকাশ করা যায়। কাজটা হাতে কলমে করা যাক--

আমরা মুখটাকে বসাতে চাই বাক্সের উপরের দিকে। দুটোকেই গ্রাফ কাগজে বসিয়ে নিই (Fig 12)। আমরা চাইছি যেন $(2,0)$ যায় $(1,1)$-এ আর $(0,2)$ যায় $(-2,1)$-এ। তার মানে canonical basis-টাকে $\left(\frac{1}{2}, \frac{1}{2}\right)$ আর $\left(-1, \frac{1}{2}\right)$-এ পাঠাতে চাইছি। সুতরাং আমাদের linear transformation-টার matrix হল

$$A = \begin{bmatrix} \frac{1}{2} & -1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Fig 11

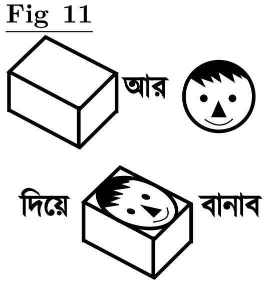


Fig 12

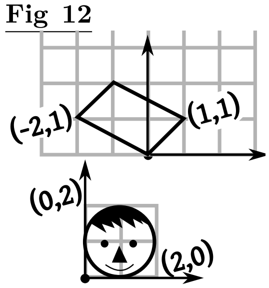
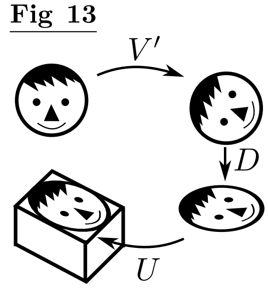


Fig 13



এর SVD হল[3]

$$A = \underbrace{\begin{bmatrix} 0.957092 & 0.2897841 \\ -0.2897841 & 0.957092 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 1.151388 & 0 \\ 0 & 0.651388 \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} 0.2897841 & -0.957092 \\ 0.957092 & 0.2897841 \end{bmatrix}}_{V'}.$$

খুবই বিচ্ছিরি লাগছে হয়তো। কিন্তু অধিকাংশ বাস্তব সমস্যাতেই এইরকম সংখ্যা এসে উপস্থিত হয়। এবার orthogonal matrix দুটো থেকে ঘোরানোর পরিমাণটা বুঝতে হবে। মনে রেখো counterclockwise দিকে (মানে ঘড়ির কাঁটার বিপরীতে) θ কোণে ঘোরালে matrix-টা হয়

$$G(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

সুতরাং যদি U-কে $G(\alpha)$ আর V'-কে $G(\beta)$ বলে ভাবতে চাই, তবে দরকার--

$$\cos\alpha = 0.957092 \quad , \quad \sin\alpha = -0.2897841,$$
$$\cos\beta = 0.2897841 \quad , \quad \sin\beta = 0.957092.$$

এ থেকেই পেয়ে যাচ্ছি $\alpha = -16.8°$, $\beta = 73.2°$. এবার ধাপগুলো বুঝে নাও (Fig 13)--প্রথমে $73.2°$ কোণে ঘড়ির কাঁটার উল্টোদিকে ঘোরাচ্ছি, তারপর x-axis বরাবর 1.151388 পরিমাণ প্রসারণ এবং y-axis বরাবর 0.651388 পরিমাণ সংকোচন করছি, এবং সব শেষে $16.8°$ কোণে ঘড়ির কাঁটার দিকে ঘোরাচ্ছি।
এই প্রয়োগটা অবশ্য নিছক মজা করে দিলাম। SVD-র আসল প্রয়োগ অন্যত্র। এবার সেটা আলোচনা করি।

64.3 Low rank approximation

অনেক সময়ে আমাদের কাজ হয় একটা matrix-কে approximate করা। হয়তো মূল matrix-টা খুব জটিল কিছু, আমরা চাই তার কাছাকাছি অপেক্ষাকৃত সহজতর কোনো matrix খুঁজে পেতে। Matrix-এর আবার "সহজ-জটিল" কী? আর দুটো matrix "কাছাকাছি" মানেই বা কী? এ দুই প্রশ্নের উত্তর নানাভাবে দেওয়া যায়, আমরা একভাবে দিচ্ছি-- যে matrix-এর rank যত বেশী, তাকে তত জটিল বলব। এরকম বলার পিছনে কারণটা এই--মনে করো একটা 100×200 সাইজের matrix আছে। সেটা তোমাকে whatsapp করে পাঠাতে হবে। তবে মোট $20,000$-টা সংখ্যা লিখতে হবে। এবার ধরো কেউ বলে দিল যে, matrix-টার rank হচ্ছে 10. তার মানে ওর 200-টা column থাকলে কী হয়, ওরা সকলেই খালি 10-টা column-এর linear combination. ধরো এই 10-টা column-এর নাম দিলাম $\vec{v}_1, ..., \vec{v}_{10}$. এরা প্রত্যেকেই $\mathbb{R}^{100}$-এর সদস্য, তাই এদের মধ্যে মোট $10 \times 1000 = 1,000$-টা সংখ্যা আছে। এবার 200-টা column-এর প্রত্যেকেই এই দশজনের linear combination. তাই সেই linear combination-এর coefficient-গুলো পাঠালেই যথেষ্ট, মানে আরো $200 \times 10 = 2,000$-টা সংখ্যা। তার মানে সব মিলিয়ে পাঠাতে হচ্ছে $1,000 + 2,000 = 3,000$-টা সংখ্যা। কুড়ি হাজারের জায়গায় মোটে তিন হাজার! যদি rank-টা r হত, তবে একই কায়দায় তোমাকে পাঠাতে হত $300r$-টা সংখ্যা। দেখতেই পাচ্ছ, r যত কমছে, ততই কম পরিশ্রমে পুরো matrix-এর যাবতীয় তথ্য পাঠানো যাচ্ছে।
এবার দ্বিতীয় প্রশ্নে আসি। দুটো matrix-কে কখন "কাছাকাছি" বলা যাবে? এর একটা উত্তর হল যখন $\sqrt{\sum_{ij}(a_{ij} - b_{ij})^2}$ ছোটো হবে। এইটাকে আমরা সংক্ষেপে লিখব $\|A - B\|$.
এই দুটো ধারণা ব্যবহার করে সহজ matrix দিয়ে জটিল matrix-কে approximate করার চেষ্টা থেকেই নীচের সংজ্ঞাটার জন্ম--

DEFINITION: Best rank-k approximation

Let $A_{m\times n}$ be any real matrix. Let $k \leq r(A)$. Then $B_{m\times n}$ is called a **best rank-k approximation** of A if $r(B) = k$ and for all matrices $C_{m\times n}$ with $r(C) = k$, we have

$$\|A - B\| \leq \|A - C\|.$$

[3] এটা কম্পিউটার দিয়ে বার করেছি।

এইটা পাওয়া যেতে পারে SVD ব্যবহার করে--

> **THEOREM**
>
> Let $A_{m\times n}$ be any real matrix with SVD $A = USV'$. Let $k \leq r(A)$. Let S_k be the same as S, but with all but the first k singular values replaced by zeroes. Then $A_k = US_kV'$ is a best rank-k approximation of A.

ইংরেজীটা কি খটোমটো লাগছে? তাহলে সহজে বোঝার জন্য একটা উদাহরণ দেখা যাক।

Example 7: এই matrix-টার একটা best rank-1 approximation দাও--

$$A = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

SOLUTION: দেখাই যাচ্ছে যে matrix-টাকে SVD হিসেবেই দেওয়া আছে। মাঝের diagonal matrix-টার প্রথম diagonal entry-টা রাখলেই best rank-1 approximation পেয়ে যাব, মানে 2-এর জায়গায় 0 বসিয়ে দেব--

$$\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & -3 \end{bmatrix}.$$

■

এই theorem-টা আমরা এখানে প্রমাণ করব না। কিন্তু এর ব্যবহারের কিছু ধারণা দেব।

64.3.1 Image compression

Fig 14-এ একটা সাদা-কালো[4] ফোটো রয়েছে।

যদি এ বইয়ের নিম্নমানের ছাপার কারণে ফোটোটা বুঝতে অসুবিধা হয়, তবে পাশের QR code-টা স্ক্যান করে নিলে ছবিটা তোমার মোবাইলের পর্দায় দেখতে পাবে।

`https://arnabc74.github.io/linalg/pixel.html`

টাইল্ বসানো মেঝে দেখেছ? এই ফোটোটাও আসলে সেরকম, অনেকগুলো ছোটো ছোটো চৌকো টালী পাশাপাশি বসিয়ে এটা তৈরী। টালীগুলোর পোশাকি নাম হল pixel. খুব ছোটো ছোটো বলে খালি চোখে ওদের আলাদা করে বোঝা যাচ্ছে না। বড় করে দেখলে ওদের অস্তিত্ব চমৎকার মালুম হবে (যেমন ফোটোটার একটা ছোটো অংশকে বড় করে পেয়েছি Fig 15)। এই pixel-গুলো সব একেকটা একরঙা টালী। রংটা একেবারে কালো থেকে একেবারে সাদা পর্যন্ত হতে পারো। কতটা কালো বা সাদা, সেটা একটা সংখ্যা দিয়ে প্রকাশ করা হয়, যাকে বলা যেতে পারে pixel-টার ঔজ্জ্বল্য। টালীগুলো যেহেতু একটা matrix-এর মত করে সাজানো, তাই তাদের ঔজ্জ্বল্যগুলো মিলে একটা matrix হবে। Fig 14-এর বেলায় এই matrix-টার

[4] এ বইটা সাদা-কালো বলে রঙীন ছবির উদাহরণ দিতে পারছি না। কিন্তু রঙীনের বেলাতেও একই কাজ করা যায়।

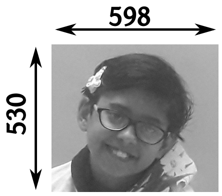


Fig 14



Fig 15



Fig 16



Fig 17

সাইজ হল 530×598. ধরো এর নাম দিলাম A. এর মধ্যে আছে $530 \times 598 = 316940$-খানা সংখ্যা। এইবার যদি আমি এর rank-20 approximation করি, তবে পাব Fig 16. এই ছবিতেও pixel-সংখ্যা একই, কিন্তু যেহেতু rank মাত্র 20, তাই প্রতিটা pixel-এর ঔজ্জ্বল্য আলাদা করে বলতে হবে না। খালি $20 \times (530 + 598) = 22560$-খানা সংখ্যা বললেই হবে, যেটা আগের 7.1% মাত্র! অবশ্যই তুমি আপত্তি তুলতে পারো যে, Fig 16 তো Fig 14-এর মত অত স্পষ্ট নয়। ঠিকই, হাজার হোক, তুমি একটা approximation করেছ। কিন্তু যদি তুমি rank-টা আরেকটু বাড়াতে, ধরো 50, তবে Fig 17-এর মত যথেষ্ট ভালো একটা approximation পেতে। এখানে তোমার 17.8% মাত্র সংখ্যা লেগেছে!

64.3.2 Neural network

এই যে একটু আগে একটা matrix-কে Whatsapp করে পাঠানোর গল্প বললাম, এই ব্যাপারটা বহু জায়গাতেই কাজে লাগে। ইন্টারনেটের মাধ্যমে যখনই জিনিসপত্র এদিকওদিক পাঠাতে হয়, তখনই এই প্রশ্নটা ওঠে--খানিকটা না হয় approximate-ই হল, তাও অল্প পরিশ্রমে একটা বড় জিনিসকে কী করে পাঠানো যায়? এর একটা নতুন প্রয়োগক্ষেত্র দেখা দিয়েছে neural network-এ। এই জিনিসটার নাম আজকাল কাগজে, টিভিতে, ওয়েবসাইটে, বিজ্ঞাপনে আকছার শোনা যায়। কিন্তু ব্যাপারটা হয়তো তোমার জানা নেই। অতি সংক্ষেপে বলি। Neural network-রা হল একধরণের function, ধরো $f : \mathbb{R}^m \to \mathbb{R}^n$. এই function-গুলোর ফর্মুলা অতিশয় জটিল হয়ে থাকে, কিন্তু এমনভাবে তৈরী যাতে একটা কম্পিউটার খুব দ্রুত সেগুলোকে বার করতে পারে। যদি $\vec{y} = f(\vec{x})$ হয়, তবে প্রতিটা y_j বার করার পিছনে প্রতিটা x_i-এরই ভূমিকা থাকতে পারে। সেইটা বোঝানোর জন্য Fig 18-এর মত ছবি আঁকা হয়। এখানে কালো কালো বিন্দুগুলোকে বলে একেকটা **node**. অংকের দৃষ্টিভঙ্গী থেকে প্রত্যেকটা node মানে একটা variable, যার value একটা সংখ্যা। লক্ষ্য করো, node-গুলো দুটো column-এ সার বেঁধে আছে। এই সারিগুলোকে বলে একেকটা **layer**. বাঁদিকের layer-টা হল input layer. এই layer-এর variable-গুলো হল $x_1, ..., x_m$. ডানদিকের layer-টা হল output layer, যার variable-গুলো হল $y_1, ..., y_n$. ছবিটা খুব ছোটো m, n-এর জন্য এঁকেছি, বাস্তবে কিন্তু এই m আর n খুবই বড় হতে পারে। তখন ছবিটা লাইনে লাইনে একেবারে কালো হয়ে ওঠার উপক্রম হবে (Fig 19)।

এই function-টার পেটের ভিতরে অনেক কাণ্ডকারখানা করা হয়, যার জন্য node-এর সংখ্যা বেশী হয়ে গেলে মুশকিল হয়। তাই ভিতরে অনেকসময়েই অল্পসংখ্যক node-ওয়ালা একটা layer দেওয়া থাকে (Fig 20)। এবার মনে করো প্রথম layer থেকে দ্বিতীয় layer-এ যাওয়ার function-টা হল g এবং দ্বিতীয় থেকে তৃতীয় layer-এ যাওয়ার function হল h. তবে দুটো মিলিয়ে প্রথম থেকে তৃতীয় layer-এ যাওয়ার function-টা দাঁড়াচ্ছে $f = h \circ g$. যদি মাঝখানে আরো অনেক অনেক layer বসাই, তবে আমাদের f-টা অনেকগুলো function-এর composition হয়ে যাবে। এটাই হল deep learning-এর মূল কথা। "Deep", অর্থাৎ কিনা প্রচুরসংখ্যক layer আছে। অবশ্য মূল ব্যাপারটা বোঝার জন্য Fig 20-এর মত একটা

Fig 18

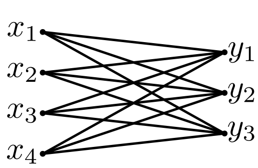


Fig 19

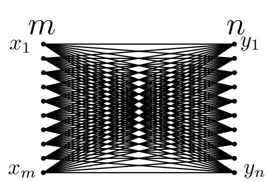
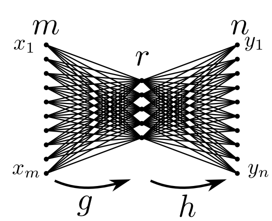


Fig 20



মধ্যবর্তী layer-ই যথেষ্ট।
যদি g আর h দুজনেই linear transformation হয়, তবে ওদেরকে matrix দিয়ে প্রকাশ করা যাবে, ধরো $B_{r\times m}$ আর $C_{n\times r}$. তবে f-ও linear transformation হবে, যার matrix হবে

$$A_{n\times m} = C_{n\times r}B_{r\times m}.$$

যখন আমরা একটা neural network তৈরী করি, তখন খালি input আর output layer দুটো এবং A-টা কী চাই সে বিষয়ে একটা ধারণা মাত্র দেওয়া থাকে। এই A-টা খুব বড় এবং জটিল কিছু হয়। আমরা পরিশ্রম লাঘব করার জন্য মাঝের ছোটো layer-টা ঢোকাই, যাতে r-খানা node আছে। এবার আমরা এমনভাবে B আর C নিতে চাই, যাতে CB যতটা সম্ভব A-র কাছাকাছি থাকে। বুঝতেই পারছ যে, $r(CB) \leq r(B) \leq r$. বিপরীতপক্ষে, r-এর চেয়ে কম rank-বিশিষ্ট যেকোনো matrix-কেই CB আকারে লেখা যায় (rank factorisation)। তার মানে CB যেকোনো matrix হতে পারে, যার rank $\leq r$. আমরা এই শর্তসাপেক্ষে A-কে approximate করার চেষ্টা করছি। সেই জন্যই low rank approximation-এর প্রশ্ন উঠছে, এবং SVD কাজে লাগছে।

64.4 Existence

আমরা এতক্ষণ SVD-র existence প্রমাণ করি নি। প্রমাণটা আসলে spectral decomposition-এর একটা সহজ প্রয়োগ। তার জন্য দুটো জিনিস লক্ষ্য করলেই হবে। ধরো $A = USV'$ হল SVD. তবে--

- এক, $A'A = (USV')'USV' = VSU'USV' = VS'SV'$, যেহেতু U-টা orthogonal হওয়ায় $U'U = I$. দেখাই যাচ্ছে যে, এটা হল $A'A$-র একটা spectral decomposition. মনে রেখো যে, $S'S$ হল একটা diagonal matrix, যার diagonal-এ শুরুতে আছে σ_i^2-রা, এবং শেষে হয়তো কিছু শূন্য আছে।
- দুই, $A = USV'$-কে লিখতে পারি $AV = US$. যদি $r(A) = r$ হয়, তবে দুপাশের প্রথম r-খানা column-কে সমান করে পাচ্ছি--
$$A\vec{v}_i = \sigma_i\vec{u}_i, \quad i = 1, ..., r. \qquad (*)$$

এই দুটো জিনিস মিলিয়ে কীভাবে আমরা যেকোনো matrix-এর জন্য SVD-র existence প্রমাণ করে ফেলব, দেখো। ধরো তুমি আমাকে একটা matrix দিয়েছ $A_{m\times n}$. আমি এটা হাতে পাওয়া মাত্র $A'A$ বার করে ফেলব। সেটা একটা real symmetric matrix. অতএব তার spectral decomposition সম্ভব। যেকোনো একটা spectral decomposition নেব, ধরো $A'A = PDP'$. এখানে D হল একটা diagonal matrix, যার diagonal-এ আছে $A'A$-র eigenvalue-রা। এদিকে $A'A$ হল একটা nnd matrix. তাই eigenvalue-গুলো সবই ≥ 0. আমরা এদের non-negative square root নেব, এবং সেগুলোকেই বড় থেকে ছোটো সাজিয়ে S বানাব। যখন আমরা বড় থেকে ছোটো সাজাচ্ছি, তখন P-এর column-গুলোকেও একইভাবে আগেপরে করে নিয়ে V তৈরী হবে। একটা উদাহরণ দেখাই। ধরো D-টা হল $diag(2,3,0)$ আর P-এর column-গুলো আছে এইরকম--

$$P = \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \vec{p}_3 \end{bmatrix}.$$

আমরা D-এর diagonal entry-গুলো সাজিয়ে পাব $S = diag(3,2,0)$. প্রথম দুটো entry-কে আগেপরে করেছি। সেই একই কাজ P-এর column-গুলোর উপরেও করে পাব--

$$V = \begin{bmatrix} \vec{p}_2 & \vec{p}_1 & \vec{p}_3 \end{bmatrix}.$$

তাহলে V আর S-এর হিল্লে হল। এবার U-এর পালা। ওর প্রথম r-টা column নেব এইভাবে--

$$\vec{u}_i = \frac{1}{\sigma_i}A\vec{v}_i, \quad i = 1, ..., r.$$

এইভাবে নেবার মতলবটা আমাদের মাথায় এসেছে (*)-টা দেখে। পড়ে আছে খালি বাকি $n-r$-খানা $\vec{u}_i$ নেওয়া। তার জন্য প্রথমে লক্ষ্য করো যে, $\{\vec{u}_1, ..., \vec{u}_r\}$ একটা orthonormal set হচ্ছে। নীচের অংকে সেটা তোমাকে দেখাতে বলেছি।

Exercise 17: দেখাও যে, $\{\vec{u}_1, ..., \vec{u}_r\}$ একটা orthonormal set হচ্ছে।

HINT:
$\|\vec{u}_i\|^2 = \vec{u}_i'\vec{u} = \frac{1}{\sigma_i^2}\vec{v}_i'A'A\vec{v}_i = 1$, কারণ $A'A\vec{v}_i = \sigma_i^2\vec{v}_i$. একইভাবে দেখাও যে, $i \neq j$ হলে $\vec{u}_i'\vec{u}_j = 0$ হবে। ■

যেকোনো orthonormal set-কেই একটা ONB-তে extend করা যায়। সেটাই করো, ব্যস্ বাকি $\vec{u}_i$-গুলোও পেয়ে গেলে। এদেরকে column হিসেবে নিয়ে U বানিয়ে নাও। আমি দাবী করছি যে, এরা মিলেই একটা SVD দেবে। কেন দেখা যাক--

- প্রথমতঃ, U আর V দুজনেই যে orthogonal matrix সে তো দেখতেই পাচ্ছ। ওদের ওভাবেই বানানো হয়েছে।
- দ্বিতীয়তঃ, $A = USV'$ হচ্ছে, কারণ সেভাবেই $\vec{u}_1, ..., \vec{u}_r$ নেওয়া হয়েছিল। এর মধ্যে $\vec{u}_{r+1}, ..., \vec{u}_n$-এর কোনো ভূমিকা নেই।

এই যুক্তিতে এক ঢিলে দুই পাখী মারা হল, existence-ও দেখানো হল, আবার হাতেকলমে SVD বানানোর একটা পথও পাওয়া গেল। এবং সেই সঙ্গে এটাও দেখতে পাচ্ছ যে, U আর V মোটেই unique নয়, যদিও S-টা unique.

64.5 Field-টা যদি $\mathbb{C}$ হয়

এতক্ষণ আমরা $\mathbb{R}$-এর উপর কাজ করছিলাম। কিন্তু field-টা যদি $\mathbb{C}$ হয়, তবেও এ সব কিছুই করা যেতে পারে। খালি তখন transpose-এর বদলে conjugate transpose নিতে হবে, এবং orthogonal-এর জায়গায় unitary. ফলে SVD-র সংজ্ঞাটা দাঁড়াবে এইরকম--

SVD (over $\mathbb{C}$)

Any complex $A_{m\times n}$ can be factorised as

$$A = USV^*$$

for some unitary $U_{m\times m}$ and $V_{n\times n}$ and diagonal $S_{m\times n}$ such that for some integer $r \geq 0$, the first r diagonal entries of S positive real numbers in non-increasing order, the rest being zeroes.

লক্ষ্য করো, এখানেও কিন্তু singular value-রা positive real number-ই আছে!
$\mathbb{C}$-এর বেলাতেও best rank-k approximation-এর ব্যাপারটা হবে। খালি দুটো matrix-এর দূরত্ব বার করতে হবে এইভাবে--

$$\|A - B\| = \sqrt{\sum_{ij} |a_{ij} - b_{ij}|^2}.$$

এখানে $|a_{ij} - b_{ij}|$ মানে complex number-এর modulus বোঝাচ্ছি।

Answers

1. যেহেতু $P^* = P^{-1}$, তাই A এবং T পরস্পরের সঙ্গে similar. সুতরাং ওদের eigenvalue-গুলো একই। এদিকে T যেহেতু triangular, তাই ওর eigenvalue-গুলোই হল ওর diagonal entry. **2.** যেহেতু $P' = P^{-1}$, তাই $AP = PD$. যদি P-এর column-গুলোকে $\vec{p}_1, ..., \vec{p}_n$ বলি, তবে AP এবং PD-র i-th column দুটোকে সমান করে পাচ্ছি $A\vec{p}_i = \lambda_i\vec{p}_i$, যেখানে $D = \text{diag}(\lambda_1, ..., \lambda_n)$. যেহেতু P-টা nonsingular, তাই $\vec{p}_i \neq \vec{0}$. সুতরাং λ_i-গুলো A-র eigenvalue এবং $\vec{p}_i$-গুলো তার eigenvector হচ্ছে। **3.** এখানে $P'AP = D$, যেখানে $P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{-2}{3} \end{bmatrix}$ এবং $D = \text{diag}(-2, 4, 1)$. **4.** যদি $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ নিই, তবে $x^2 + y^2 + z^2 = 1$ এবং

$\vec{x}'M\vec{x} = x^2 + 2y^2 + 3z^2$. ঠিক যেন আমি বাজারে গেছি কিছু টাকা নিয়ে (মোট টাকার পরিমাণটাকে আমি 1 বলছি), তা দিয়ে তিন ধরণের চাল কিনতে পারি, যদি খালি প্রথম চালটাই কিনতাম, তবে সেটা 1 কেজি কেনা যেত, খালি দ্বিতীয় চালটা কিনলে 2 কেজি, খালি তৃতীয় চালটা কিনলে 3 কেজি। আমি মিলিয়ে মিশিয়ে কিনছি, x^2 টাকার প্রথম চাল, y^2 টাকার দ্বিতীয় চাল, এবং অবশিষ্ট z^2 টাকার তৃতীয় চাল। তবে মোট কেনা হচ্ছে $x^2 + 2y^2 + 3z^2$ কেজি চাল। আমরা এই পরিমাণটাকেই maximum করতে চাই। বুঝতেই পারছ, এখানে প্রথম আর দ্বিতীয় চালদুটোকে কিনে টাকা নষ্ট করার মানে হয় না, পুরো টাকাটাই তৃতীয় চালের পিছনে ঢালা উচিত, কারণ সেটাই সবচেয়ে সস্তা। মানে $x = y = 0$ এবং $z = \pm 1$ নেওয়া উচিত। সেক্ষেত্রে maximum-টা হবে 3. **6.** (A), কারণ eigenvalue-গুলো সবাই negative, অথচ ওদের গুণফল হল determinant-টা, যেটা দেখা যাচ্ছে positive. সেটা হতে পারে একমাত্র তখনই যখন জোড়সংখ্যক negative সংখ্যা গুণ করা হবে, কারণ মাইনাসগুলো জোড়ায় জোড়ায় কাটাকাটি হয়ে প্লাস হয়ে যাবে। **7.** Triangular matrix, তাই diagonal entry-গুলোই হল eigenvalue. এদের মধ্যে positive আর negative দুইই আছে, তাই matrix-টা indefinite হবে। **8.** $\frac{x^2}{4} - \frac{y^2}{3} = 1$. Hyperbola. **9.** **(1)** $\frac{x^2}{4} - \frac{y^2}{3} = 1$, hyperbola **(2)** $\frac{x^2}{4} + \frac{y^2}{3} = 1$, ellipse **(3)** $x^2 + y^2 = 5$, circle **(4)** $y = 6x^2$, parabola **(5)** কোনোটাই নয়। **10.** Rotation matrix-টা দেখতে হয় $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$-এর মত। এদের determinant হয় 1. সুতরাং $\begin{bmatrix} R & \vec{0} \\ \vec{0}' & 1 \end{bmatrix}$-এর determinant-ও 1 হবে। তাই নতুন matrix-এর determinant-টা শুরুর matrix-এর determinant-এর সমান হবে।
11. **(1)** $m = 3, n = 4, r = 2$. **(2)** অসম্ভব, কারণ diagonal নয়। **(3)** $m = n = 2, r = 0$.
(4) অসম্ভব, diagonal-এ negative থাকা চলবে না। **(5)** অসম্ভব, 0-টা শেষে থাকা উচিত ছিল।
12. শেষে একটা identity matrix কল্পনা করলেই SVD পেয়ে যাবে--

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

সুতরাং singular value-গুলো হল 2 আর 1. এদের প্রথমটার জন্য একটা left singular vector হল $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ (বাঁদিকের matrix-টার প্রথম column) আর right singular vector হল $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ (ডানদিকের matrix-টার প্রথম row-টা খাড়া করে লেখা)। দ্বিতীয়টার জন্য একটা left singular vector হল $\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ (বাঁদিকের matrix-টার দ্বিতীয় column) আর right singular vector হল $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ (ডানদিকের matrix-টার দ্বিতীয় row)। **13.** PIP'-এর দুপ্রান্তে orthogonal matrix আছে, মাঝে আছে diagonal matrix. সেই degree matrix-এর diagonal বরাবর খালি positive সংখ্যা আছে। যেহেতু সংখ্যাগুলো সবই 1, তাই বলা যেতে পারে যে, ওরা ছোটো থেকে বড় সাজানো (মানে কোথাও ছোটোর পরে বড় আসেনি)। ব্যস্, SVD হতে এর চেয়ে বেশী আর কী লাগে?
14. $A'A = VS'U'USV' = VS'SV'$. এখানে $S'S$ একটা diagonal matrix. সুতরাং $V(S'S)V'$ হল $A'A$-এর একটা spectral decomposition. তাই $A'A$-র eigenvalue-রা হবে $S'S$-এর diagonal entry-গুলো। এই diagonal entry-গুলো S-এর diagonal entry-গুলো square. **15.** হ্যাঁ, কারণ $AA' = USV'VS'U' = USS'U'$. **16.** $AA' = [14]$, তাই A-র singular value একটাই, $\sqrt{14}$. আবার $B'B = [17]$. সুতরাং B-এর singular value-ও একটাই, $\sqrt{17}$.
17. $\langle \vec{u}_i, \vec{u}_j \rangle = \vec{u}_i'\vec{u}_j = \frac{1}{\sigma_i\sigma_j}\vec{v}_i'A'A\vec{v}_j = \frac{\sigma_j}{\sigma_i}\vec{v}_i'\vec{v}_j = 0$, কারণ $A'A\vec{v}_j = \sigma_j^2\vec{v}_j$ এবং $\vec{v}_i'\vec{v}_j = 0$.

Chapter X
Jordan/Rational Canonical Form

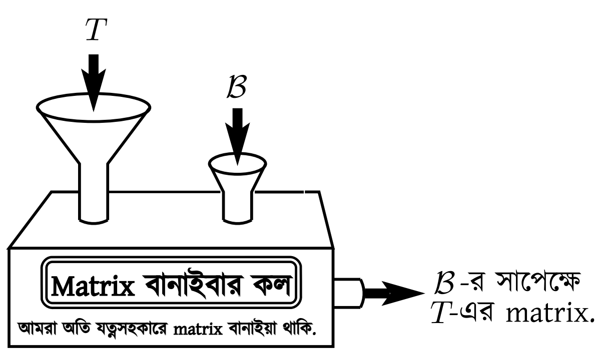
DAY 65
Minimal polynomial

65.1 এই অধ্যায়ের উদ্দেশ্য

এই অধ্যায়ে আমরা যা আলোচনা করতে চলেছি, সেটা linear algebra প্রয়োগের জন্য ভীষণ দরকারী বলা যায় না। কিন্তু একটা গুরুত্বপূর্ণ তাত্ত্বিক প্রশ্নের উত্তর এখানে আমরা শিখব। প্রশ্নটার উল্লেখ আমরা ইতিপূর্বেই করেছি (Fig 1)। আরেকবার মনে করিয়ে দিই।

Linear algebra শেখার সময়ে matrix নিয়ে আমরা খামোখা মাথা ঘামাই কেন? বিভিন্ন কারণ আছে, তার মধ্যে সবচেয়ে প্রধান কারণ হল এই যে, matrix দিয়ে linear transformation-দের প্রকাশ করা যায়। ঘুরিয়ে বললে আমাদের আসল মাথাব্যথাটা linear transformation-দের নিয়ে, matrix-রা নিমিত্ত মাত্র। অনেকটা যেভাবে লোকে ডাক্তারের ফোন নম্বর মুখস্থ করে রাখে, সেটা ঠিক সংখ্যাপ্রীতির কারণেই নয়, সময়ে-অসময়ে ডাক্তারের প্রয়োজনের কথা ভেবেই! একই ডাক্তারের যদি একাধিক ফোন নম্বর থাকে, তবে তাদের মধ্যে কোনটা ব্যবহার করছি, সেটা যেমন গুরুত্বপূর্ণ নয়, ডাক্তারের সঙ্গে যোগাযোগ করতে পারলেই হল, তেমনি একই linear transformation-এর জন্য যদি একাধিক matrix থাকে, তবে তাদের মধ্যে আমরা বাছবিচার করতে চাই না। কিন্তু একই linear transformation-এর জন্য কি বিভিন্ন matrix পাওয়া সম্ভব? উত্তর হল--হ্যাঁ, যদি basis বদলাও। আলোচনার সুবিধার জন্য ধরো আমাদের linear transformation-এর domain আর codomain একই। তার নাম দিলাম V, যেটা ধরো $\mathbb{C}$-এর উপরে একটা finite dimensional vector space. মানে আমাদের linear transformation-টা হল $T : V \to V$. তবে V-এর যেকোনো একটা basis নিলে T-এর জন্য একটা matrix পাবে। যদি linear transformation-টা একই থাকে, তবে বিভিন্ন basis-এর জন্য তার matrix-গুলোকে বলি পরস্পরের সঙ্গে **similar**. দুটো matrix দেওয়া থাকলে আমরা স্বাভাবিকভাবেই জানতে চাই ওরা পরস্পরের সঙ্গে similar কিনা। তার জন্য কোনো সহজে পরীক্ষা করে দেখা যায় এরকম necessary and sufficient শর্ত খুঁজে পাওয়া সহজ ছিল না। Necessary শর্ত প্রচুর পাওয়া গেছিল, যেমন trace, determinant, eigenvalue, characteristic polynomial, ইত্যাদি। কিন্তু এরা কেউই sufficient হয় নি। এমনকি এদের সবাইকে একসঙ্গে মিলিয়েও একটা sufficient শর্ত বানানো যায় নি। অর্থাৎ এমন এক জোড়া matrix বার করা গেছে যাদের trace, determinant, rank, nullity, eigenvalue,

Fig 1



characteristic polynomial ইত্যাদি সবই একই, অথচ তাও যারা পরস্পরের সঙ্গে similar নয়। এই sufficient শর্ত খুঁজে পাওয়ার সমস্যাটা গণিতজ্ঞদের বহুদিন ভুগিয়েছিল। অবশেষে তার যে সব সমাধান বেরিয়েছে, এই অধ্যায়ে আমরা সেরকম দুটো আলোচনা করব। এখানে সেই পরিচিত উপমাটা ফের খাটে। একই fraction-কে আমরা নানাভাবে লিখতে পারি, যেমন $\frac{2}{4}$, $\frac{5}{10}$ এবং $\frac{7}{14}$ আসলে একই fraction. দুটো fraction সমান কিনা সেটা দেখার পথ হল দুটোকেই পুরো কাটাকাটি করে লিখে দেখা একই জিনিস আসে কিনা। দুটো matrix পরস্পরের সঙ্গে similar কিনা, তার পরীক্ষাও অনেকটা সেরকম। Matrix দুটোকে একধরণের "কাটাকাটি" করে অতি সহজ রূপে পরিণত করা যায়। এরকম দুটো সহজ রূপ আমরা আলোচনা করব। একটার নাম **Jordan canonical form** বা **JCF**. তার জন্য field-টাকে $\mathbb{C}$ হতে হয়। আর অন্যটার নাম **rational canonical form** বা **RCF**. সেটা যেকোনো field-এর জন্যই কাজ করে। কিন্তু সেই প্রসঙ্গে যাওয়ার আগে আমরা আরো একটা necessary condition শিখব (যেটাও sufficient নয়)।

65.2 Minimal polynomial

আমরা Cayley-Hamilton theorem থেকে দেখেছিলাম যে, যদি কোনো square matrix A-র characteristic polynomial হয় $p(\lambda)$, তবে $p(A) = O$ হয়। যেমন নীচের উদাহরণে।

Example 1: $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ হলে তার characteristic polynomial বার করো। যদি তাকে $p(\lambda)$ নাম দিই, তবে দেখাও যে, $p(A) = O$ হবে।

SOLUTION:এখানে characteristic polynomial হল--

$$p(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 4 \end{vmatrix} = (\lambda - 1)(\lambda - 4) - 6 = \lambda^2 - 5\lambda - 2.$$

যখন $p(A)$ বার করব ওই শেষের -2-টার জায়গায় $-2I$ নিতে হবে--

$$p(A) = A^2 - 5A - 2I = \cdots = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

■

কিন্তু characteristic polynomial ছাড়াও এমন অন্য বহু polynomial সম্ভব যাদের লাগালে A-টা শূন্য matrix-এর পরিণত হয়। এই সব polynomial-দের একটা বিশেষ নাম আছে--

DEFINITION: Annihilating polynomial

An **annihilating polynomial** for a square matrix A is any polynomial $f(\lambda)$ such that $f(A) = O$.

ইংরাজিতে "annihilate" মানে "নির্মূল করা"।

Example 2: যদি $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ নিই, তবে নীচের polynomial তিনটের মধ্যে কে বা কারা A-র annihilating polynomial হবে?

$$f(\lambda) = \lambda - 1, \quad g(\lambda) = (\lambda - 1)^2 \text{ এবং } h(\lambda) = (\lambda - 1)^3.$$

SOLUTION: এখানে $f(A) = A - I \neq$). তাই $f(\lambda)$ মোটেই A-র annihilating polynomial হচ্ছে না।
যখন $g(A)$ বার করবে তখন দুভাবে করতে পারো--

- এক, $g(\lambda) = \lambda^2 - 2\lambda + 1$ লিখে নিলে, $g(A) = A^2 - 2A + I = O$ পাবে।
- দুই, সরাসরি $g(A) = (A - I)^2 = O$ বার করতে পারো।

তার মানে $g(\lambda)$ অবশ্যই A-র একটা annihilating polynomial হচ্ছে।
এবার $h(A)$-র বেলায় আর নতুন করে অংক কষতে হবে না, কারণ $h(A) = (A - I)^3 = g(A)(A - I)$. যেহেতু $g(A) = O$, তাই $h(A) = O$ না হয়ে যায় না! অতএব $h(\lambda)$-ও A-র একটা annihilating polynomial. ■

এই অংকটা থেকেই বুঝতে পারছ যে, যদি $f(\lambda)$ কোনো A-র একটা annihilating polynomial হয়, তবে $f(\lambda)$-এর যেকোনো multiple-ই (মানে $f(\lambda)$ দিয়ে ভাগ যাওয়া যেকোনো polynomial-ই) A-র annihilating polynomial হবে।

Exercise 1: ধরো $B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. সবচেয়ে ছোটো কোন্ $k \in \mathbb{N}$-এর জন্য $p(\lambda) = (\lambda - 2)^k$ এই B-এর একটা annihilating polynomial হবে? ■

Cayley-Hamilton theorem থেকে আমরা বলতে পারি যে, characteristic polynomial হল একটা annihilating polynomial. এই annihilating polynomial-দের মধ্যে একটাই polynomial থাকবে যার degree সবচেয়ে কম এবং যেটা monic[1]. একে বলে ওই matrix-এর **minimal polynomial.**

> DEFINITION: **Minimal polynomial**
>
> If A is a square matrix, then a polynomial $p(\lambda)$ is called its **minimal polynomial** if it is monic, is an annihilating polynomial, and has the minimum degree among all non-zero annihilating polynomials of A.

এই সংজ্ঞার মধ্যে কয়েকটা বিষয় লক্ষণীয়--

- প্রথমতঃ, আমরা minimal polynomial বার করার সময়ে খালি non-zero annihilating polynomial-দের মধ্যে degree-টা minimum রাখার চেষ্টা করছি। Zero polynomial-ও কিন্তু একটা annihilating polynomial (বস্তুতঃ zero polynomial যেকোনো square matrix-এরই annihilating polynomial)। কিন্তু zero polynomial-এর degree ধরা হয় $-\infty$. তাই minimum বার করার সময়ে ওকে হিসেবে নিলে, ওটাই সব matrix-এর minimal polynomial হয়ে যেত।
- দ্বিতীয়তঃ, যেকোনো square matrix-এরই একটা minimal polynomial যে থাকবেই, তার কোনো গ্যারান্টি কিন্তু এই সংজ্ঞার মধ্যে দেওয়া নেই। তবে গ্যারান্টি সত্যিই দেওয়া যায়, শুধু একটা নয়, দু-দুটো। যেকোনো square matrix-এর জন্যই minimal polynomial পাওয়া যাবে, এবং সেটা আবার unique-ও হবে! তার প্রমাণ আমরা শীঘ্রই করব।

প্রমাণের প্রসঙ্গে যাবার আগে একটা ঠকানো অংক দিই, যেটা থেকে বোঝা যাবে তুমি সংজ্ঞাটা ভালো করে বুঝেছ কিনা।

Exercise 2: বলো তো 9×9 সাইজের শূন্য matrix-এর minimal polynomial কী হবে?

(A) $m(\lambda) \equiv 1$ (B) $m(\lambda) \equiv \lambda$ (C) $m(\lambda) \equiv x$ (D) এদের কোনোটাই নয়।

[1] একটা polynomial-কে monic বলে যদি তার সবচেয়ে বড় power-এর coefficient-টা 1 হয়। যেমন $\lambda^2 - 5\lambda + 6$ হল monic, কিন্তু $-\lambda^2 - 5\lambda + 6$ মোটেই monic নয়।

■

Exercise 3: এমন একটা matrix দিতে পারো, যার minimal polynomial হল $m(\lambda) \equiv 1$? ■

এই অংকটা থেকে যা শিখলে, সেটাকেই গুছিয়ে লিখলে নীচের theorem-টা হয়--

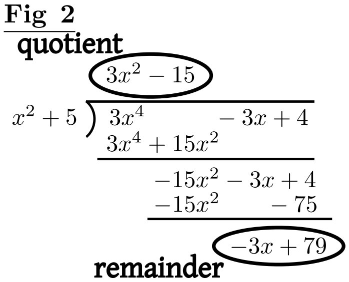
> **THEOREM**
>
> The minimal polynomial of a matrix cannot be a constant polynomial.

এর মধ্যে existence অংশটার কারণ বোঝাটা তেমন শক্ত নয়। তোমাকে যদি অনেকগুলো positive integer দিই, তবে তুমি সর্বদাই তাদের মধ্যে থেকে সবচেয়ে ছোটো একজনকে পাবে, তাই না? সাবধান, সবচেয়ে বড় একজন নাও পেতে পারো (যদি আমি তোমাকে infinitely many সংখ্যা দিয়ে থাকি), কিন্তু সবচেয়ে ছোটো একটা পাবেই, কারণ ওদের মধ্যে যেকোনো একটা সংখ্যা নাও, ধরো 5. ব্যস্, অমনি তুমি $1, 2, 3, 4$ আর 5 এই কয়টা সংখ্যার মধ্যে খুঁজে দেখো কোন্‌টা কোন্‌টা তোমার set-এ আছে। এই যে ব্যাপারটা, এটার একটা গালভরা নাম আছে--**well-ordering principle**. এবার এই ব্যাপারটা তুমি খাটাও যাবতীয় annihilating polynomial-এর degree-র উপরে। তবে এমন একটা annihilating polynomial তুমি পাবেই, যার degree-টা ওদের মধ্যে minimum. সুতরাং সংজ্ঞার দুটো শর্ত পূরণ হল, annihilating এবং minimum degree. এবার monic শর্ত পূরণ করতে হবে। ধরো আমরা minimum degree annihilating polynomial পেয়েছি $5x^2 - 3x + 4$. এটা monic নয়, কারণ তার জন্য x^2-এর coefficient-টা 1 হওয়া দরকার ছিল, কিন্তু এখানে সেটা 5. কিন্তু তাতে কী? 5 দিয়ে পুরোটাকে ভাগ করে দাও, তাতেও জিনিসটা একটা annihilating polynomial-ই থাকবে, এবং ওর degree-ও বদলে যাবে না। সুতরাং তুমি পেয়ে যাবে $x^2 - \frac{3}{5}x + \frac{4}{5}$, যেটা minimal polynomial হবার সব শর্তই পালন করে। এই যুক্তিতে আমরা চুপিচুপি Cayley-Hamilton theorem লাগিয়েছি, সেটা টের পেয়েছ কি? ওখান থেকেই এই গ্যারান্টিটা পেয়েছি যে, অন্ততঃ একটা annihilating polynomial আছে।

এবার দেখাব যে, minimal polynomial-টা unique হয়। তার প্রমাণটা অপেক্ষাকৃত লম্বা। ধাপে ধাপে এগোব। প্রথমে তোমাকে ছোটোবেলায় শেখা ভাগের অংক একটু মনে করিয়ে দিই। যদি তোমাকে বলি 17-কে 3 দিয়ে ভাগ করলে remainder (ভাগশেষ) কী হবে, তবে উত্তর হল 2, কারণ $17 = 3 \times 5 + 2$. ওই 5-টা এখানে quotient (ভাগফল)। নিশ্চয়ই মনে আছে যে, যদি $n > 0$ দিয়ে ভাগ করো, তবে ভাগশেষটা হবে $0, 1, ..., n-1$-এর মধ্যে কিছু একটা। একই ব্যাপার করা যায় polynomial-দের নিয়েও--একটা polynomial-কে আরেকটা polynomial দিয়ে ভাগ করে ভাগফল এবং ভাগশেষ বার করা যায়। সেটাও স্কুলে থাকতেই শিখেছিলে (Fig 2)। এক্ষেত্রে remainder-এর degree-টা divisor-এর (মানে ভাজকের) degree-র চেয়ে কম হতে বাধ্য। এই জিনিসটাকে গুছিয়ে লিখলে বলা যায়--

> যদি $p(x)$ আর $d(x)$ যেকোনো দুটো polynomial হয় (যেখানে $d(x)$ হল nonzero), তবে এমন দুটো polynomial পাবে $q(x)$ আর $r(x)$ যাতে
>
> $$p(x) = d(x)q(x) + r(x)$$

Fig 2



> হয়, এবং $\deg(r(x)) < \deg(d(x))$ হয়। এই $q(x)$ এবং $r(x)$ দুজনেই unique. এদের মধ্যে $q(x)$-কে বলে **quotient** (ভাগফল) আর $r(x)$-কে বলে **remainder** (ভাগশেষ)। চাইলে আরেকবার Fig 2 দেখে নিতে পারো।

এইটা আমাদের এক্ষুণি কাজে লাগবে। আর একটা notation-ও একটু মনে করিয়ে দিই, যেটা আমরা স্কুলে থাকতে সাধারণতঃ শিখি না। যদি $a(x)$ আর $b(x)$ দুটো polynomial হয়, তবে $a(x)|b(x)$ লেখা মানে "$a(x)$ divides $b(x)$", অর্থাৎ সোজা বাংলায় "$a(x)$ দিয়ে $b(x)$-কে ভাগ যায়", অর্থাৎ $b(x)$-কে $a(x)$ দিয়ে ভাগ করলে remainder হবে 0. যদি $a(x)|b(x)$ এবং $b(x)|a(x)$ দুটোই একই সঙ্গে হয়, তবে $a(x) = \alpha b(x)$ হবে, যেখানে α কোনো একটা nonzero সংখ্যা।
Minimal polynomial-এর uniqueness প্রমাণ করার পরবর্তী ধাপ হল নীচের theorem-টা।

THEOREM

Minimal polynomial must divide all annihilating polynomials.

Proof:

> Let $m(\lambda)$ be the minimal polynomial of a square matrix A, and let $p(\lambda)$ be any annihilating polynomial of A.

দেখাতে হবে যে, $m(\lambda)|p(\lambda)$. তার জন্য $p(\lambda)$-কে $m(\lambda)$ দিয়ে ভাগ করে শুরু করা যাক। ধরো quotient এল $q(\lambda)$ আর remainder এল $r(\lambda)$, অর্থাৎ--

> We know that
> $$p(\lambda) = m(\lambda)q(\lambda) + r(\lambda)$$
> for some polynomials $q(\lambda)$ and $r(\lambda)$ with $\deg(r(\lambda)) < \deg(m(\lambda))$.
> Then
> $$\begin{aligned} & p(A) = m(A)q(A) + r(A) \\ \textit{or} \quad & O = Oq(A) + r(A) \quad [\because m(\lambda), p(\lambda) \text{ both annihilate } A] \\ \textit{or} \quad & r(A) = O. \end{aligned}$$

সুতরাং $r(\lambda)$-ও একটা annihilating polynomial হচ্ছে।

> But since $m(\lambda)$ has minimum degree among all nonzero annihilating polynomials, we must have $r(\lambda) \equiv 0$, completing the proof.

[Q.E.D]

এই theorem-টা ব্যবহার করে আমরা দেখাতে পারব যে, যেকোনো square matrix-এর জন্যই minimal polynomial হয় unique.

THEOREM

Every square matrix has a unique minimal polynomial.

Proof: এই theorem-টার মধ্যে existence আর uniqueness দুটোই ধরা আছে। Existence-এর যুক্তিটা আমরা আগেই আলোচনা করেছি। এখানে সংক্ষেপে লিখে দিই।

> Existence: Let D be the degree of all nonzero annihilating polynomials of a matrix. By Cayley-Hamilton theorem we know that the characteristic polynomial is a non-zero annihilating polynomial.
>
> So D is a nonempty subset of $\mathbb{N}$.
>
> Hence, by the well-ordering property of $\mathbb{N}$, D has minimum element, d, say.
>
> So there is some annihilating polynomial of degree d. Dividing it by its leading coefficient we get a minimal polynomial.

এখানে leading coefficient মানে সবচেয়ে বড় power-এর coefficient.

> Uniqueness: Let $m_1(x)$ and $m_2(x)$ both be minimal polynomials for the same square matrix.

তার মানে দুজনেই annihilating polynomial-ও বটে। আমরা জানি minimal polynomial দিয়ে annihilating polynomial-দের ভাগ যায়। তাই--

> Then, by the last theorem, $m_1(x)|m_2(x)$ and $m_2(x)|m_1(x)$.

এবার লক্ষ্য করো যে, যদি দুটো polynomial-এর দুজনকে দিয়েই দুজনে ভাগ যায়, তবে ওরা পরস্পরের nonzero multiple হতে বাধ্য, যেমন $2\lambda^2 - 2$ আর $3\lambda^2 - 3$.

> Hence $m_2(x) = \alpha m_1(x)$ for some non-zero scalar, α.
> But since both the polynomials are monic, hence $\alpha = 1$, completing the proof.

[Q.E.D]

65.3 Minimal polynomial বার করা

আমরা minimal polynomial-এর existence আর uniqueness প্রমাণ করলাম। এর পরের প্রশ্ন হল একটা square matrix দেওয়া থাকলে কী করে তার minimal polynomial বার করা যাবে। এই অধ্যায়ে পরে যখন আমরা JCF বার করা শিখব, তখন তা থেকে minimal polynomial-ও অনায়াসে বেরিয়ে যাবে। কিন্তু খালি minimal polynomial বার করার জন্য JCF বার করতে যাওয়া অনেক ক্ষেত্রেই "মশা-মারতে-কামান-দাগা" হয়ে যায়। ছোটোখাটো matrix-এর বেলায় তার চেয়ে অনেক সহজ পথেই কাজ হয়, একটু এদিক-ওদিক ঢুঁ মেরে। সেই কায়দাটা বলি।

আমরা দেখেছি যে, minimal polynomial-টা যেকোনো annihilating polynomial-কেই ভাগ করে। আরও দেখেছি যে, characteristic polynomial হল একটা annihilating polynomial. সুতরাং--

> **Theorem**
>
> The minimal polynomial must divide the characteristic polynomial.

সুতরাং minimal polynomial বার করার শুরুতেই characteristic polynomial বার করে নিলে সুবিধা হবে। Characteristic polynomial-টা বার করা সহজ, স্রেফ গায়ের জোরে একটা determinant বার করার ব্যাপার। তারপর minimal polynomial-এর সন্ধানটা characteristic polynomial-টার factor-গুলোর (মানে উৎপাদকগুলোর) মধ্যে করলেই হবে।

Example 3: ধরো কোনো real matrix-এর characteristic polynomial হল λ^2+1. তার minimal polynomial কী হবে?

SOLUTION: আমরা জানি যে, minimal polynomial কখনো একটা constant polynomial হতে পারে না। এদিকে λ^2+1-এর খালি একটাই non-constant, monic factor আছে, যেটা ও নিজেই। তাই minimal polynomial-ও হবে ওই λ^2+1-টাই। ■

এই যুক্তিতে বুঝতে পারছ যে, characteristic polynomial যদি irreducible হয় (মানে উৎপাদকে বিশ্লেষণ না করা যায়), তবে minimal polynomial-টা characteristic polynomial-এর সমান হবে। কিন্তু ধরো characteristic polynomial-টা যদি হয় $(\lambda-1)(\lambda-2)$, তবে? এর যেসব non-constant, monic factor আছে, তারা হল $(\lambda-1)$, $(\lambda-2)$ এবং $(\lambda-1)(\lambda-2)$ নিজে। এদের মধ্যে কোনটা minimal polynomial হবে? এর উত্তর আসবে minimal polynomial-এর আরেকটা ধর্ম থেকে--

> characteristic polynomial-টা যেখানে যেখানে শূন্য হবে, minimal polynomial-টাও সে সব জায়গায় শূন্য হতে বাধ্য!

দিব্যি জিনিসটা, গুছিয়ে লিখে প্রমাণ করে রাখার মত--

> THEOREM
>
> Let $p(\lambda)$ be the characteristic polynomial and $m(\lambda)$ be the minimal polynomial of a square matrix, A over a field F. Let $\alpha \in F$ be such that $p(\alpha)=0$. Then $m(\alpha)=0$.

Proof:

> Since $p(\alpha)=0$, hence α must be an eigenvalue of A. So $m(\alpha)$ must be an eigenvalue of $m(A)$.

এটা এল eigenvalue-র একটা ধর্ম থেকে, যেটা আমরা সপ্তম অধ্যায়ে শিখেছিলাম--কোনো matrix-এর একটা polynomial নিলে eigenvalue-গুলোর উপরেও একই polynomial চাপে। যেমন A-র একটা eigenvalue যদি λ হয়, তবে $\lambda^2-3\lambda+1$ হবে A^2-3A+I-এর একটা eigenvalue.

> But $m(A)=O$.

কারণ minimal polynomial-রা annihilating polynomial-ও বটে।

> Hence $m(\alpha)=0$, as required.

এখানে আমরা এই তথ্যটা ব্যবহার করলাম যে, zero matrix-এর সব eigenvalue-ই শূন্য হয়। **[Q.E.D]**

একটা ছোট্টো ঠকানো প্রশ্ন--

Exercise 4: একটা matrix-এর সব eigenvalue-ই শূন্য। তার মানে matrix-টা কি zero matrix হতে বাধ্য? ■

এইবার আমাদের আগের অংকটাকে গুছিয়ে আক্রমণ করা যাক।

Example 4: যদি characteristic polynomial হয় $(\lambda-1)(\lambda-2)$, তবে minimal polynomial কী হবে?

SOLUTION: এখানে characteristic polynomial-টা দুই জায়গায় শূন্য হচ্ছে 1-এ আর 2-এ। সুতরাং minimal polynomial-কেও ওই দুই জায়গায় শূন্য হতে হবে। তার মানে $m(\lambda)$-র মধ্যেও $(\lambda-1)(\lambda-2)$ থাকবেই। এদিকে $m(\lambda)|p(\lambda)$ হতে হবে। আবার $m(\lambda)$-কে monic-ও হতে হবে। তাই $m(\lambda)=(\lambda-1)(\lambda-2)$ হওয়া ভিন্ন গত্যন্তর নেই। ■

এই যুক্তিটাকে আরেকটু এগিয়ে নিয়ে যাওয়া যায়।

Example 5: ধরো কোনো real matrix-এর জন্য characteristic polynomial হল $p(x)=(x^2+1)(x-3)$. তবে $m(x)$ কী হবে?

SOLUTION: এখানে $m(x)$-এর মধ্যে $(x-3)$ যে থাকবে সেটা বোঝাই যাচ্ছে। কিন্তু x^2+1-টা কি থাকবে? আগের অংকের যুক্তি এখানে সরাসরি খাটানো যাচ্ছে না, কারণ x^2+1 শূন্য হবে $x=\pm i$-তে, যারা কেউ $\mathbb{R}$-এর মধ্যে নেই। এখানে তাই আমরা একটু কৌশল করব। যেহেতু যেকোনো real সংখ্যাকেই complex সংখ্যা বলেও ভাবা যায়, তাই আমরা শুরুতে কিছুক্ষণ field-টাকে $\mathbb{R}$-এর বদলে $\mathbb{C}$ কল্পনা করে নেব। তবে $p(x)=(x-i)(x+i)(x-3)$ লিখতে পারব। সুতরাং এবার আগের অংকের যুক্তিতে $\mathbb{C}$-এর উপরে minimal polynomial-টা $(x-i)(x+i)(x-3)=(x^2+1)(x-3)$ হবে। কিন্তু $m(x)=(x^2+1)(x-3)$-টা আসলে দিব্যি $\mathbb{R}$-এর উপরেও একটা polynomial. দেখাই যাচ্ছে যে, $m(A)=O$ এবং $\mathbb{R}$-এর উপরে এর চেয়ে কম degree-র কোনো annihilating polynomial সম্ভব নয়, কারণ--

> যদি $\mathbb{R}$-এর উপরে এর চেয়ে কম degree-র কোনো annihilating polynomial থাকত, তবে সেটা $\mathbb{C}$-এর উপরেও একটা annihilating polynomial-ও হত। কিন্তু সেটা অসম্ভব, যেহেতু $\mathbb{C}$-এর উপরে $m(x)$ হল minimal polynomial.

তাই $\mathbb{R}$-এর উপরেও minimal polynomial হবে $m(x)=(x^2+1)(x-3)$. ■

এতক্ষণ পর্যন্ত আলোচিত সবগুলো characteristic polynomial-এর মধ্যেই প্রতিটা factor-এর power ছিল 1 করে। যদি আরো বড় power থাকে, তবে কী হবে? সেই নিয়েই নীচের অংকটা।

Example 6: যদি $p(x)=(x^2+2)^2(x-1)^3(x-4)^5$ হয়, তবে $m(x)=?$

SOLUTION: এখানে $m(x)$-এর মধ্যে অন্ততঃ একটা x^2+2, একটা $x-1$ আর একটা $x-4$ চাই। কিন্তু একটাতেই যে মন উঠবে, এমন নাও হতে পারে। তাই এখানে এই চেহারার প্রত্যেকটা polynomial-কেই পরীক্ষা করে দেখতে হবে--

$$m(x)=(x^2+2)^r(x-1)^s(x-4)^t,\ \text{যেখানে } 1\leq r\leq 2,\ 1\leq s\leq 3 \text{ এবং } 1\leq t\leq 5.$$

প্রতিটা r,s,t-র জন্যই $m(A)$ বার করে করে দেখতে হবে সবচেয়ে কম degree-তে কখন শূন্য হয়। এর কোনো শর্টকাট হয় না? উত্তর হল--হয়, যদি matrix-টা JCF বা RCF তোমার জানা থাকে। এই অধ্যায়ের বাকি অংশে যখন JCF আর RCF-এর কথা শিখবে, তখন দেখবে যে, উপরের প্রতিটা r,s,t-র জন্যই এমন A সম্ভব, যার characteristic polynomial হল $p(x)$, আর minimal polynomial হল $(x^2+2)^r(x-1)^s(x-4)^t$. ■

Exercise 5: নীচের প্রতিক্ষেত্রে একটা করে matrix-এর characteristic polynomial দেওয়া আছে। তোমাকে বলতে হবে minimal polynomial-গুলো কী কী হবে। যদি কোনো ক্ষেত্রে একাধিক সম্ভাবনা থাকে, তবে সবগুলো সম্ভাবনা বলো।

1. $p(x)=(x-1)(x-2)(x-3)$

2. $p(x) = (x-1)^2(x-2)(x-3)$
3. $p(x) = (x^2+1)(x^2+2)$
4. $p(x) = x^3$

■

65.4 Necessary, কিন্তু sufficient নয়

আমরা শুরুতেই বলেছিলাম যে, minimal polynomial-রা similar হবার জন্য necessary শর্ত, কিন্তু sufficient নয়। এর মানে হল-- দুটো matrix যদি পরস্পরের সঙ্গে similar হয়, তবে তাদের minimal polynomial-ও সমান হতে বাধ্য। কিন্তু দুটো matrix-এর minimal polynomial সমান হলেই জোর দিয়ে বলা যাবে না যে, ওরা similar হবেই। এই দুটো জিনিস গুছিয়ে লিখে একে একে প্রমাণ করা যাক। প্রথমে necessary অংশটা--

THEOREM

If A and B are two similar matrices, then they have the same minimal polynomial.

Proof: আমরা প্রথমে দেখাব যে, A আর B-এর সব annihilating polynomial-রাই একই, মানে যে polynomial-এ A ঘায়েল হয়, সেই একই polynomial-এ B-ও ঘায়েল হবে।

> Since A and B are similar, hence $A = RBR^{-1}$ for some nonsingular matrix, R.
> Let $p(\lambda)$ be an annihilating polynomial for A. Then
> $$p(A) = R\,p(B)\,R^{-1}.$$

কারণ $A^2 = RBR^{-1}RBR^{-1} = RB^2R^{-1}$. একইভাবে যেকোনো $n \in \mathbb{N}$-এর জন্যই $A^n = RB^nR^{-1}$. সুতরাং যেকোনো polynomial-এর জন্যই এই similarity-টা বজায় থাকবে।

> Thus $R\,p(B)\,R^{-1} = O$, or $p(B) = O$.
> So all the annihilating polynomials of A are also annihilating polynomials for B.
> By interchanging the roles of A and B, all the annihilating polynomials of B are also annihilating polynomials of A.
> Hence A, B have the same set of annihilating polynomials.

যেহেতু A আর B-এর annihilating polynomial-দের set-টা একই, তাই এই set-এর মধ্যে সবচেয়ে কম degree-র monic polynomial-টাই দুজনেরই minimal polynomial হবে।

> So the monic annihilating polynomial with minimum degree is also the same for both A and B.
> Hence A and B have the same minimal polynomial, as required.

[Q.E.D]

এবার একটা counterexample দিয়ে দেখাব যে, minimal polynomial-রা similar হবার জন্য sufficient নয়।

Example 7: এমন দুটো একই সাইজের matrix দাও, যারা পরস্পরের সঙ্গে similar নয়, কিন্তু যাদের minimal polynomial একই।

SOLUTION: এইরকম একজোড়া matrix আমি বলে দিচ্ছি। কিন্তু কী করে এদের জোগাড় করলাম, সেটা এখনই বোঝাতে পারব না। এই অধ্যায়ে পরে যখন JCF শিখবে, তখন সেটা বুঝতে পারবে।

Matrix দুটো হল--

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ এবং } B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

একটু চিন্তা করলেই বুঝবে যে, $A^2 = B^2 = O$. যেহেতু $A, B \neq O$, তাই দুইক্ষেত্রেই minimal polynomial হল $p(\lambda) = \lambda^2$.

কিন্তু A আর B যে similar নয়, সে তো বোঝাই যায়, কারণ ওদের rank অসমান, $r(A) = 2$ এবং $r(B) = 1$. ■

DAY 66
Jordan Canonical Form (part 1)

66.1 ব্যাপারটা কী?

এক কথায় বলতে গেলে Jordan Canonical Form (JCF) মানে হল complex field-এর উপরে একটা বিশেষ চেহারার square matrix। ঠিক যেমন diagonal matrix বা upper triangular matrix, সেইরকম। চেহারাটা খুবই সহজ, চোখে দেখলেই চেনা যায়। শুধু তাই নয়, ওই চেহারার দুটো matrix পরস্পরের সঙ্গে similar কিনা, সেটাও চোখে দেখেই বোঝা যায়। যদি A যেকোনো square matrix হয়, তবে তুমি A-কে ওই চেহারার এমন একটা matrix-এ পরিণত করতে পারবে, যেটা ওই A-র সঙ্গে similar. সুতরাং যদি দুটো square matrix দিই, A আর B, এবং পরীক্ষা করে দেখতে বলি ওরা similar কিনা, তবে তুমি দুজনকেই JCF-এ পরিণত করবে, এবং তারপর চোখে দেখেই উত্তরটা পেয়ে যাবে।

এই হল সব খুঁটিনাটি বাদ দিয়ে মোদ্দা গল্পটা। এবার খুঁটিনাটিগুলো বলব। এর জন্য প্রথমে লাগবে এই সংজ্ঞাটা--

Definition: Jordan block

For $n \in \mathbb{N}$ and $\lambda \in \mathbb{C}$ the **Jordan block** $J_n(\lambda)$ denotes the square matrix

$$J_n(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}_{n\times n}.$$

Here all the blank entries are zeroes.

লক্ষ্য করো, এটা একটা upper triangular matrix. এর diagonal entry-গুলো সবাই λ.

Example 8: এই তিনটে Jordan block-কে বড় করে লেখো, $J_3(4)$, $J_1(2+i)$ এবং $J_2(0)$.

Solution: $J_3(4) = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$. $J_1(2+i) = \begin{bmatrix} 2+i \end{bmatrix}$. $J_2(0) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. ■

Exercise 6: $J_3(7+2i)$-কে বড় করে লেখো। ■

Exercise 7: $J_k(a) - J_k(b) = ?$ ■

Exercise 8: $J_k(6) - 2I$-কে $J_m(\lambda)$ আকারে লিখলে m আর λ কত? ■

Exercise 9: $\begin{bmatrix} J_2(3) & O \\ O & J_2(3) \end{bmatrix}$ আর $J_4(3)$ কি একই matrix? এখানে O বলতে 2×2 সাইজের শূন্য matrix বুঝিয়েছি। ■

যদি এরকম এক বা একাধিক Jordan block-কে কোণাকুণি জুড়ে একটা block diagonal matrix বানাও, তবে সেরকম matrix-কে বলে একটা JCF. যেহেতু প্রতিটা diagonal block-ই upper triangular, তাই পুরো matrix-টাও তাই।

Example 9: নীচের matrix-টা কী একটা JCF? এর মধ্যে কী কী Jordan block আছে?

$$\begin{bmatrix} 2 & 1 & \\ & 2 & \\ & & 3 \end{bmatrix}.$$

এখানে ফাঁকা জায়গাগুলোতে সব শূন্য আছে।

Solution:হ্যাঁ, এটা একটা JCF. যদি এটাকে এভাবে ভাবো--

$$\begin{bmatrix} \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix} & \\ & \begin{bmatrix} 3 \end{bmatrix} \end{bmatrix},$$

তবেই বুঝবে যে, এর মধ্যে দুটো Jordan block আছে, $J_2(2)$ আর $J_1(3)$. ■

Example 10: নীচের matrix-টা কি JCF চেহারায় আছে?

$$\begin{bmatrix} 2 & 1 & \\ & 3 & \\ & & 3 \end{bmatrix}.$$

SOLUTION:না, এটা মোটেই JCF নয়। 2-এর ব্লকটা শেষ হবার আগেই 3 শুরু হয়ে গেছে। ■

Exercise 10: একটা JCF লেখো, যার মধ্যে দুটো $J_2(3)$ আর একটা $J_3(2)$ আছে। ■

Exercise 11: ধরো একটা diagonal matrix আছে, যার diagonal entry-গুলো হল $1, 2, 3, 4$. এটা কি JCF-এ আছে? থাকলে এর Jordan block-গুলো কী কী? ■

66.2 JCF-দের existence এবং uniqueness

এবার আমরা JCF-দের মূল ধর্ম দুটোর কথা বলব। প্রথমতঃ, যেকোনো complex, square matrix-এর জন্যই এমন একটা JCF পাওয়া যায়, যেটা সেই matrix-এর সঙ্গে similar.

> THEOREM
>
> For any complex, square matrix there is a matrix in JCF similar to it.

এই গেল existence-এর কথা। এবার uniqueness-এর কথা বলি। আমরা বলেছিলাম যে, দুটো matrix যদি JCF-এ থাকে, তবে ওরা similar কিনা, সেটা চোখে দেখেই বলা যায়। তার কারণ এরকম দুটো matrix পরস্পরের সঙ্গে similar হবে if and only if ওদের মধ্যে একই সংখ্যক একই রকমের Jordan block থাকবে, খালি বড়জোর কে আগে কে পরে, সেটা আলাদা হতে পারে।

Example 11: নীচের JCF দুটো কি পরস্পরের সঙ্গে similar?

$$\begin{bmatrix} J_2(3) & & & \\ & J_1(2) & & \\ & & J_2(3) & \\ & & & J_3(i) \end{bmatrix} \quad \begin{bmatrix} J_3(i) & & & \\ & J_1(2) & & \\ & & J_2(3) & \\ & & & J_2(3) \end{bmatrix}.$$

SOLUTION: হ্যাঁ, কারণ প্রথমটায় আছে পরপর এই ব্লকগুলো $J_2(3)$, $J_1(2)$, $J_2(3)$ এবং $J_3(i)$. অন্যটার বেলায় পরপর রয়েছে $J_3(i)$, $J_1(2)$, $J_2(3)$ এবং ফের আরেকটা $J_2(3)$. তার মানে ব্লকগুলো একই, খালি ক্রমের পার্থক্য। ■

Example 12: ধরো দুটো JCF আছে। প্রথমটায় পরপর আছে $J_2(3)$, $J_1(2)$, $J_2(3)$ এবং $J_3(i)$. অন্যটার বেলায় পরপর রয়েছে $J_3(i)$, $J_1(3)$, $J_2(2)$ এবং $J_2(3)$ এই JCF দুটো কি পরস্পরের সঙ্গে similar?

SOLUTION:না, এরা similar হতেই পারে না। ■

এই যে ব্যাপারটাকেই গুছিয়ে লিখলে এই theorem-টা পাবে।

> THEOREM
>
> Two complex, square matrices in JCF are similar to each other if and only if they contain the same numbers of same Jordan blocks (possibly in a different order).

আবার সেই fraction-এর উপমাটা মনে করা যাক। কোনো fraction-এর যেমন একটা reduced form হয়, তেমনি যেকোনো complex, square matrix-এর একটা JCF হয়। বাড়তি প্যাঁচ খালি এই যে, reduced form-টা হত unique. এখানে Jordan block-গুলোর ক্রমটা নিয়ে খেলা করা যেতে পারে। তাই আমরা বলি যে, একটা complex, square matrix-এর জন্য JCF-টা হয় unique up to the order of the Jordan blocks.
Fraction-এর বেলায় reduced form-এ পৌঁছনোর একটা কায়দা আমরা জানতাম--কাটাকাটি করা। JCF-এর বেলাতেও সেরকম একটা পথ আছে। এবার সেটাই শিখব।

66.3 একটা matrix-এর JCF বার করা

ধরো তোমাকে একটা complex, square matrix দিয়ে বললাম তার একটা JCF বার করতে। তার একটা কায়দা আছে। ঠিক সহজ কায়দা নয় সেটা, তবে সম্পূর্ণ যান্ত্রিক, কোনো বুদ্ধি খাটাতে হয় না, গড়গড় করে গেলেই উত্তর বেরোতে বাধ্য। কায়দাটার জন্য তোমাকে দুটো জিনিস করতে পারতে হবে--

- এক, matrix-টার eigenvalue বার করে ফেলা, তাদের algebraic multiplicity-সুদ্ধ।
- আর দুই, কোনো matrix-এর nullity বার করতে পারা। যেহেতু rank+nullity=number of columns হয়, তাই rank বার করতে পারলেও হবে। যদি B কোনো matrix হয়, তবে তার nullity আমরা বোঝাব $\nu(B)$ দিয়ে।

পুরো প্রক্রিয়াটা বোঝাচ্ছি একটা উদাহরণ দিয়ে।

- তোমার প্রথম কাজ হবে matrix-টার যাবতীয় eigenvalue বার করে ফেলা, algebraic multiplicity-সুদ্ধ।

  উদাহরণস্বরূপে--ধরো ঠিক দুটো eigenvalue আছে $1+i$ আর 0. প্রথমটার algebraic multiplicity ধরো 10, আর দ্বিতীয়টার 3.

- এর পরে একটা কাজ করতে হবে, প্রতিটা eigenvalue-র জন্য আলাদা করে। যে eigenvalue-টা নিয়ে কাজ করছ, সেটাকে যদি λ বলি, তবে তোমাকে $B=A-\lambda I$-এর nullity বার করতে হবে। এবং পাশাপাশি ততগুলো বিন্দু আঁকতে হবে।

  আমাদের উদাহরণে-- ধরো আমরা প্রথমে $\lambda=1+i$ নিয়ে কাজ করব। শুরুতে $B=A-(1+i)I$-এর nullity বার করব। ধরো $\nu(B)=4$ পেলে। অমনি কাগজে চারটে বিন্দু পাশাপাশি আঁকো-- • • • •
  .

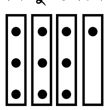
- Algebraic multiplicity-এর সমানসংখ্যক বিন্দু কি হয়ে গেছে? যদি হয়ে গিয়ে থাকে, তবে এই eigenvalue নিয়ে কাজ শেষ। নইলে এবার বার করো $\nu(B^2)$. এটা সর্বদাই $\nu(B)$-এর চেয়ে বেশী হবে। যত বেশী, ততগুলো বিন্দু আঁকো আগের বিন্দুগুলোর নীচে, বাঁদিক থেকে।

আমাদের উদাহরণে--algebraic multiplicity ছিল 10, কিন্তু বিন্দু পেয়েছি 4-টা। যেহেতু $4 < 10$, তাই $\nu(B^2)$ বার করতে হবে। ধরো $\nu(B^2) = 7$ পেলে। আগে ছিল 4. তাই নতুন বিন্দু লাগবে তিনটে, সেগুলোকে আগের বিন্দুগুলোর নীচে নীচে লিখব, বাঁদিক থেকে --
• • • •
• • • ।

- পরের ধাপগুলোও একইরকম। শুরুতে দেখে নিতে হবে algebraic multiplicity-র সমান সংখ্যক বিন্দু হল কিনা। না হলে $\nu(B), \nu(B^2), \nu(B^3), \ldots$ এইভাবে বার করে যেতে হবে। তার জন্য যত যত বাড়তি বিন্দু লাগে, সেগুলো লিখবে তার পর পর লাইনে, বাঁদিক থেকে।

  আমাদের উদাহরণে--সাতটা বিন্দু হয়েছিল, এদিকে algebraic multiplicity হল 10. তাই $\nu(B^3)$ লাগবে। ধরো $\nu(B^3)$ হল 10. তার মানে আরো তিনটে বিন্দু চাই। আগের বিন্দুগুলোর নীচে এঁকে দাও--
  • • • •
  • • • ।
  • • •

- যখন algebraic multiplicity-র সমান সংখ্যক বিন্দু পেয়ে যাবে, তখন দেখতে হবে কতগুলো column হয়েছে। ঠিক ততগুলোই Jordan block হবে সেই eigenvalue-র জন্য। যেই column-এ যতগুলো বিন্দু, সেই Jordan block-এর সেইটাই হবে সাইজ।

  আমাদের উদাহরণে--আমরা $\lambda = 1 + i$ নিয়ে কাজ করছিলাম। $\mathrm{am}(\lambda, A) = 10$ ছিল। আমরা 10-টা বিন্দু পেয়ে গেছি, সুতরাং এই eigenvalue নিয়ে কাজ শেষ। আমরা চারটে column পেয়েছি--
  
  ওদের সাইজ $3, 3, 3$ এবং 1. তার মানে এখানে তিনটে $J_3(1 + i)$ এবং একটা $J_1(1 + i)$ থাকবে।

- এবার একই কাজ ফের করতে হবে পরের eigenvalue-টা নিয়ে।

  আমাদের উদাহরণে-- অন্য eigenvalue-টা হল $\lambda = 0$, যার algebraic multiplicity ছিল 3. তাই এবার $B = A - 0I = A$ নেব, এবং $\nu(B)$, $\nu(B^2), \ldots$ ইত্যাদি বার করার প্রস্তুতি নেব। ধরো $\nu(B) = 3$ হয়ে গেল, মানে তিনটে বিন্দু • • • . যেহেতু algebraic multiplicity-র সমান বিন্দু পেয়ে গেছি, তাই গল্প এখানেই শেষ। তিনটে column আছে, প্রত্যেকটাতেই একটা করে বিন্দু। তার মানে তিনটে $J_1(0)$ ব্লক হবে।

- এবার সব eigenvalue-র জন্য পাওয়া Jordan block-দের পরপর সাজিয়ে দিলেই হবে। ক্রমটা গুরুত্বপূর্ণ নয় মোটেই।

  আমাদের উদাহরণে-- হাতে আছে তিনটে $J_3(1 + i)$, একটা $J_1(1 + i)$ আর তিনটে $J_1(0)$. এদের যেভাবে খুশি সাজিয়ে একটা block diagonal matrix বানালেই হল।

কায়দাটা কি সর্বদাই কাজ করবে? উত্তর হল--হ্যাঁ, করবে, কোথাও গোলমাল হবে না। যেমন ধরো
• • •
• • • • •
• • • •
-এর মত অবস্থা কখনো হবে না, যেখানে কোনো লাইনে তার আগের লাইনের চেয়ে বেশী বিন্দু হয়ে গেছে। আবার যদি algebraic multiplicity হয় 8, তবে
• • • •
• • •
• •
-এর মত অবস্থাও হবে না, যেখানে প্রথম দুটো লাইনে মোট বিন্দুর সংখ্যা < 8, অথচ শেষ লাইনটাকে হিসেবে নিলে মোট বিন্দুর সংখ্যা হয়ে যাচ্ছে > 8. কোনো একটা লাইনের শেষে ঠিক algebraic

multiplicity-র সমান সংখ্যক বিন্দুই হবে, এবং সেখানেই তুমি থামবে। সেই থামার আগে পর্যন্ত প্রত্যেক লাইনেই তুমি অন্ততঃ একটা করে বিন্দু পাবেই। এই সব কিছুই আমরা এই অধ্যায়ের শেষে গিয়ে প্রমাণ করব।
পুরো প্রক্রিয়াটা সড়গড় করে নেবার জন্য নীচের অংকটা করো।

Exercise 12: ধরো $A_{10\times 10}$-এর ঠিক তিনটে eigenvalue আছে, 5, i আর 3. এদের algebraic multiplicity যথাক্রমে 5, 3 এবং 2. বলে দিচ্ছি যে--

- $\nu(A-5I)=2,\ \nu((A-5I)^2)=4$ এবং $\nu((A-5I)^3)=5,$
- $\nu(A-iI)=2$ এবং $\nu((A-iI)^2)=3$
- $\nu(A-3I)=1.$

তবে A-র একটা JCF বার করো। ■

প্রক্রিয়াটা যান্ত্রিক হলেও খুব যে একটা সহজ এমন বলা যায় না। কারণ সব eigenvalue এবং তাদের algebraic multiplicity বার করা সহজ নাও হতে পারে। আর B, B^2, B^3 ইত্যাদিদের nullity বার করাও বেশ বিরক্তিকর। কিন্তু কিছু বিশেষ বিশেষ ক্ষেত্রে এত খাটনি করতে হয় না। এবার সেরকম কিছু উদাহরণ দেখাই। যাদের মাথায় পরীক্ষাপাশের চিন্তা, তারা এই উদাহরণগুলো ভালো করে দেখো।

66.4 কিছু সহজ ক্ষেত্র

আমরা এক্ষুণি যে বিন্দু বসানোর কায়দাটা বললাম, সেখানে প্রতিটা eigenvalue-র জন্য কয়েক লাইন ধরে বিন্দু আঁকতে হচ্ছিল। মোট বিন্দুর সংখ্যা ছিল eigenvalue-টার algebraic multiplicity আর প্রথম লাইনে বিন্দুর সংখ্যা ছিল $\nu(A-\lambda I)$, যেটা আসলে geometric multiplicity. তাই eigenvalue-টার algebraic multiplicity আর geometric multiplicity জানলেই বিন্দুগুলোর গঠন অনেকটাই বেরিয়ে যায়। বস্তুতঃ কিছু বিশেষ ক্ষেত্রে পুরো গঠনটাই বার করে ফেলা যায় স্রেফ এই দুটি multiplicity থেকেই। এরকম কয়েকটা ক্ষেত্র আলোচনা করব এবার।

66.4.1 $\mathrm{gm}(\lambda, A) = \mathrm{am}(\lambda, A)$

যদি কোনো eigenvalue-র বেলায় algebraic multiplicity=geometric multiplicity হয়, তবে সব বিন্দুই প্রথম লাইনে থাকবে, পরের লাইনের জন্য কোনো বিন্দু পড়ে থাকবে না। এরকম eigenvalue-দের যে regular eigenvalue বলে, সেটা আমরা সপ্তম অধ্যায়েই জেনেছি। এরকম eigenvalue-র জন্য সব বিন্দুই এক লাইনে থাকায় ওই eigenvalue-র জন্য algebraic multiplicity-সংখ্যক Jordan block থাকে, যাদের প্রতিটার সাইজ হয় 1×1, মানে JCF-এর মধ্যে ওই eigenvalue-র জন্য অংশটা diagonal হয়।

Example 13: একটা matrix-এর তিনটে eigenvalue, $2, 4, 0$. এদের algebraic multiplicity হল যথাক্রমে 2, 3 এবং 2. যদি এদের geometric multiplicity-ও তাই হয়, তবে একটা JCF বার করো।
SOLUTION: যেকোনো একটা diagonal matrix নিলেই হবে, যার সাইজ 7×7 এবং যার diagonal entry-গুলো মধ্যে দুটো 2, তিনটে 4 এবং দুটো 0 আছে। ক্রমটা গুরুত্বপূর্ণ নয়। ■

Exercise 13: ধরো

$$A = \begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 2 & * \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

এখানে $*$-এর জায়গায় যেকোনো complex সংখ্যা থাকতে পারে। বলে দেওয়া আছে যে, এর সব eigenvalue-র geometric multiplicity-ই 2. তবে একটা JCF বার করো।

■

যদি একটা matrix কোনো diagonal matrix-এর সঙ্গে similar হয়, তবে তার সব eigenvalue-ই হবে regular. এটা আমরা সপ্তম অধ্যায়েই দেখেছিলাম। এক্ষেত্রে সেই diagonal matrix-টাই হবে আমাদের matrix-এর একটা JCF. সপ্তম অধ্যায়ে আমরা কয়েকটা শর্ত শিখেছিলাম, যেগুলো থেকে কোনো eigenvalue-কে regular দেখানো যায়--

- যদি কোনো eigenvalue-র algebraic multiplicity-টা 1 হয়, তবে geometric multiplicity-ও 1 হবে, এবং তাই eigenvalue-টা regular হবে।

- যদি কোনো $n \times n$ সাইজের matrix-এর সব eigenvalue-এর geometric multiplicity-দের যোগফল হয় n, তবেও সবগুলো eigenvalue-ই regular হবে।

সুতরাং এইসব ক্ষেত্রেও JCF বার করার সুবিধা আছে।

66.4.2 $\mathrm{gm}(\lambda, A) = 1$

ধরো একটা eigenvalue আছে, যার algebraic multiplicity হল 3 আর geometric multiplicity হল 1. সুতরাং প্রথম লাইনে একটা বিন্দু থাকবে। আমরা জানি[2] প্রতি লাইনেই অন্ততঃ একটা বিন্দু থাকবে, এবং কোনো লাইনেই তার আগের লাইনের চেয়ে বেশী বিন্দু থাকতে পারে না। তাই প্রথম লাইনেই যদি মোটে একটা বিন্দু থাকে, তবে পরের প্রতিটা লাইনেও ঠিক একখানা করেই বিন্দু থাকতে বাধ্য। তাই বিন্দুর গঠনটা হবে $\begin{array}{c}\bullet\\\bullet\\\bullet\end{array}$. যেহেতু একটাই column হচ্ছে, তাই এরকম কোনো eigenvalue-র জন্য ঠিক একটাই Jordan block থাকবে, যার সাইজ algebraic multiplicity-এর সমান।

Example 14: একটা 5×5 সাইজের matrix-এর ঠিক একটাই eigenvalue আছে, 4. যদি সেটার geometric multiplicity হয় 1, তবে একটা JCF বার করো।

SOLUTION: যেহেতু 5×5 সাইজের matrix-এর একটাই eigenvalue, তাই algebraic multiplicity হবে 5. তাই মোট 5-খানা বিন্দু আছে, একটার উপর একটা খাড়াভাবে। সুতরাং উত্তর হবে $J_5(4)$. ■

66.4.3 $\mathrm{gm}(\lambda, A) = \mathrm{am}(\lambda, A) - 1$

যদি geometric multiplicity-টা algebraic multiplicity-এর চেয়ে ঠিক এক কম হয়, তবে বিন্দুর গঠনে ঠিক দুটো লাইন থাকবে, এবং দ্বিতীয় লাইনের জন্য খালি একটাই বিন্দু পড়ে থাকবে।

Example 15: ধরো λ হল এমন একটা eigenvalue, যার algebraic multiplicity হল 3 এবং geometric multiplicity হল 2. এই eigenvalue-টা কী কী Jordan block সরবরাহ করবে?

SOLUTION: এখানে বিন্দুর গঠন হবে $\begin{array}{ccc}\bullet&\bullet&\bullet\\\bullet&&\end{array}$. সুতরাং এখানে একটা $J_2(\lambda)$ আর দুটো $J_1(\lambda)$ থাকবে। ■

Exercise 14: একটা 9×9 সাইজের matrix-এর একমাত্র eigenvalue হল 2, যার geometric multiplicity হল 8. এই matrix-টার JCF-এ কী কী Jordan block কয়টা করে আছে? ■

Exercise 15: একটা 10×10 সাইজের complex, square matrix আছে, যার eigenvalue-রা হল 0, i এবং -2. এদের algebraic multiplicity হল যথাক্রমে 4, 2 আর 4, এবং geometric multiplicity হল যথাক্রমে 3, 2

[2] এগুলো সব প্রমাণ করব এই অধ্যায়ের শেষে গিয়ে।

আর 1. তবে এই matrix-টার JCF-এ কী কী Jordan block কয়টা করে আছে, লেখো। ■

66.4.4 3×3 matrix

এ যাবৎ যে কয়টা বিশেষ ক্ষেত্র শিখলাম, তাদের ব্যবহার করে যেকোনো 3×3 সাইজের matrix-এর বেলায় সব eigenvalue-র algebraic multiplicity এবং geometric multiplicity জানলেই JCF বার করে দেওয়া যায়। কারণ প্রতিটা eigenvalue-র জন্যই এই কেসগুলোর মধ্যে একটা হতে বাধ্য (ধরো A হল matrix-টা, আর λ হল eigenvalue-টা)--

- $\mathrm{am}(\lambda, A) = \mathrm{gm}(\lambda, A) = k$, যেখানে $k = 1, 2, 3$. সেক্ষেত্রে ওই eigenvalue-র জন্য একটাই Jordan block আসবে, $J_k(\lambda)$.
- $\mathrm{am}(\lambda, A) = k$ আর $\mathrm{gm}(\lambda, A) = 1$. যেখানে $k = 2, 3$. এক্ষেত্রে খালি একখানা Jordan block আসবে, $J_k(\lambda)$.
- $\mathrm{am}(\lambda, A) = 3$ আর $\mathrm{gm}(\lambda, A) = 2$. এক্ষেত্রে Jordan block আসবে দুটো, একখানা $J_2(\lambda)$ আর একখানা $J_1(\lambda)$.

প্রতি ক্ষেত্রে বিন্দুগুলোর গঠন কীরকম হবে, সেটা Fig 3-এ দেখিয়েছি।

Example 16: একটা 3×3 সাইজের matrix আছে, যার eigenvalue আছে খালি দুটো λ আর μ. এদের মধ্যে λ-র algebraic multiplicity হল 2, এবং geometric multiplicity হল 1. তবে একটা JCF বার করো।

SOLUTION: বোঝাই যাচ্ছে যে, μ-এর algebraic multiplicity হবে $3 - 2 = 1$ এবং geometric multiplicity-ও তাই। যেহেতু λ-র geometric multiplicity হল 1, তাই ওর জন্য খালি একটাই Jordan block হবে। সুতরাং Jordan block-গুলো হচ্ছে $J_2(\lambda)$ এবং $J_1(\mu)$. তাই একটা JCF হবে--

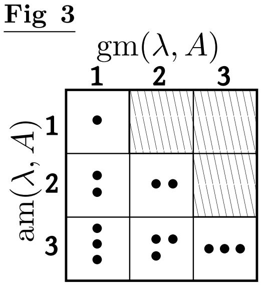
$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix}.$$

■

Example 17: Show that the matrix $A = \begin{bmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{bmatrix}$ has a Jordan canonical form. Find a Jordan canonical form of A. What are the number of distinct Jordan canonical forms of A?[1+3+1] **(2020.CC12.3h)**

SOLUTION:যেকোনো complex, square matrix-এরই JCF থাকে। সেটাই হল JCF-এর প্রধান theorem. সেই theorem-টা লিখে দিলেই প্রথম অংশের উত্তর শেষ।

Fig 3



> By a standard theorem, every complex, square matrix over $\mathbb{C}$ has a JCF. Since the given matrix is a square matrix with real entries (which may be considered as complex entries), it has a JCF.

দ্বিতীয় অংশের জন্য characteristic polynomial-টা বার করে নিই, যাতে সবগুলো eigenvalue এবং তাদের algebraic multiplicity-গুলো পাওয়া যায়।

> The characteristic polynomial of the matrix is
>
> $$-\lambda^3+7\,\lambda^2-16\,\lambda+12=-(\lambda-3)(\lambda-2)^2.$$
>
> So we have only two distinct eigenvalues, 2 and 3, with algebraic multiplicities 2 and 1, respectively.

বাধ্য হয়েই 3-এর geometric multiplicity-ও 1-ই হবে। সুতরাং ওর অবদান হবে খালি একটা $J_1(3)$.

> So $\mathrm{gm}(3,A)=1$. Hence this will produce a single $J_1(3)$.

অন্যটার বেলায় geometric multiplicity-টা বার করে দেখতে হবে--

> Now $A-2I=\begin{bmatrix}1 & 1 & -2\\ -1 & -2 & 5\\ -1 & -1 & 2\end{bmatrix}$.

দেখাই যাচ্ছে যে, এর প্রথম আর তৃতীয় row পরস্পরের negative. কিন্তু দ্বিতীয় row-টা ওদের multiple নয়। সুতরাং rank হচ্ছে 2. তাই nullity হবে $3-2=1$.

> This has rank 2.
>
> ⟦ <u>Because:</u>
>
> The 1st and 3rd rows are nonzero and multiples of each other, however the 2nd row is not their multiple.
>
> ⟧
>
> Hence nullity $3-2=1$. Thus, it contributes a single Jordan block $J_2(2)$.

তার মানে পেলাম মোট দুটো ব্লক। এদেরকে যেকোনো ক্রমে লিখলেই একটা JCF পাব--

> So one possible JCF is
>
> $$\begin{bmatrix}2 & 1 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{bmatrix}.$$

এবার তৃতীয় অংশ। এটা একটু হাস্যকর, কারণ JCF বানানোর পিছনে মূল উদ্দেশ্যটাই হল similar matrix-দের শনাক্ত করা, এবং তার জন্য ব্লকগুলোর ক্রমটাকে অগ্রাহ্য করাই কাজ। কিন্তু এখানে সেটা থেকেই প্রশ্ন বানিয়েছে। দুটো distinct জিনিসকে 2!-ভাবে সাজানো যায়। সেটাই উত্তর--

> The total number of JCF here is the number of ways the two distinct objects $J_2(2)$ and $J_1(3)$ may be ordered, ie, $2! = 2$.

■

Exercise 16: Find a JCF of the matrix $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix}$. ■

DAY 67
Jordan Canonical Form (part 2)

আজকে আমরা JCF নিয়ে আরেকটু গভীরে ঢুকব। প্রথমে একটা notation বানিয়ে রাখলে সুবিধা হবে। যদি একটা JCF-এ পরপর দুটো $J_4(2+i)$, একটা $J_2(-3)$ আর একটা $J_1(0)$ থাকে, তবে আমরা JCF-টাকে লিখব $\text{diag}(J_4(2+i), J_4(2+i), J_2(-3), J_1(0))$. এভাবে লিখে জায়গাও যেমন বাঁচবে, তেমনি কোনো JCF-এর power বার করতেও সুবিধা হবে, কারণ সেই আমরা প্রথম অধ্যায়ে partitioned matrix-দের গুণের কায়দা শিখেছিলাম, সেখান থেকেই বলা যাবে যে, এই JCF-টার 5-th power পাওয়া যাবে Jordan block-গুলোকে আলাদা করে 5-th power-এ তুললে, মানে--

$$\text{diag}(J_4(2+i), J_4(2+i), J_2(-3), J_1(0))^5 = \text{diag}(J_4(2+i)^5, J_4(2+i)^5, J_2(-3)^5, J_1(0)^5).$$

এইভাবে power করাটা শীঘ্রই আমাদের কাজ লাগবে।

67.1 JCF থেকে কী কী জানা যায়?

আমরা আগেই বলেছিলাম যে, দুটো square matrix-কে আমরা similar বলি, যদি ওরা একই linear transformation-কে প্রকাশ করে (আলাদা basis-এর সাপেক্ষে)। Matrix নিয়ে মাথা ঘামাবার অন্যতম প্রধান উদ্দেশ্য হল সেটা যে linear transformation-টাকে বোঝাচ্ছে, সেটার ধর্মগুলো বার করা। একটা matrix-এর পিছনের linear transformation-টার সম্পর্কে যা যা জানা সম্ভব, সবই তার JCF থেকে জানা যায়। তার মধ্যে রহস্য কিছু নেই, কারণ matrix-টা তার নিজের JCF-এর সঙ্গে similar, আর similar matrix-দের মধ্যে যেকোনো একজনের determinant, rank, eigenvalue ইত্যাদি বার করতে পারলেই অন্যটারগুলোও জানা হয়ে গেল! কিন্তু JCF-এর একটা বাড়তি সুবিধা আছে। ওটার দিকে এক ঝলক তাকিয়েই rank, eigenvalue, determinant, minimal polynomial ইত্যাদি বলে দেওয়া যায়, নতুন করে খাতা পেন বাগিয়ে অংক কষতে হয় না। সেটাই এবার আলোচনা করব।

67.1.1 Eigenvalue, algebraic multiplicity, characteristic polynomial

সবচেয়ে সহজ হল eigenvalue-গুলো বার করা। Diagonal বরাবর ওরা থরে থরে সাজানোই আছে। বস্তুতঃ যেকোনো triangular matrix-এর ক্ষেত্রেই এ কথা খাটে। যে eigenvalue-টা diagonal-এ যতবার এসেছে, সেটাই হল তার algebraic multiplicity.

Example 18: ধরো এই JCF-টা দিলাম--

$$\text{diag}(J_2(4), J_2(4), J_2(4), J_5(4), J_2(0), J_2(0)).$$

এর eigenvalue-গুলো কী কী, এবং তাদের algebraic multiplicity-ই বা কত?
SOLUTION: এখানে eigenvalue-গুলো হল 4 আর 0. তিনটে $J_2(4)$ আর একটা $J_5(4)$ মিলিয়ে সেটা diagonal entry হিসেবে এসেছে $3\times 2+1\times 5 = 11$ বার। সেটাই হল 4-এর algebraic multiplicity. একইভাবে 0-র algebraic

multiplicity হল $2 \times 2 = 4$. ■

Exercise 17: যে JCF-টা পাঁচটা $J_1(3)$, দুটো $J_2(1)$ আর তিনটে $J_2(3)$ দিয়ে তৈরী, তার eigenvalue-গুলো এবং তাদের algebraic multiplicity বার করো। ■

যেই একবার eigenvalue এবং algebraic multiplicity পেয়ে গেলে, অমনি পুরো matrix-টার characteristic polynomial লিখে ফেলতেও কোনো অসুবিধা নেই। কারণ, যদি eigenvalue-গুলো হয় $\lambda_1 \neq \cdots \neq \lambda_k$ আর তাদের algebraic multiplicity-গুলো হয় $m_1, ..., m_k$, তবে characteristic polynomial হবে

$$p(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}.$$

Example 19: একটা matrix-এর JCF-এ রয়েছে তিনটে $J_2(3)$ আর একটা $J_1(-3)$. এর eigenvalue-গুলো কী কী? Algebraic multiplicity-গুলো কত কত? আর characteristic polynomial-ই বা কী?

SOLUTION: এখানে diagonal-এ আছে 3 আর -3. সুতরাং ওরাই হল eigenvalue. প্রতিটা $J_2(3)$ দুটো করে 3 সরবরাহ করছে, এরকম তিনটে Jordan block আছে, মানে মোট হল $2 \times 3 = 6$-টা 3. তাই 3-এর algebraic multiplicity হল 6. ওদিকে -3 রয়েছে খালি 1-টা, তাই -3-এর algebraic multiplicity হল 1. সুতরাং characteristic polynomial হবে $(\lambda - 3)^6(\lambda + 3)$. ■

Exercise 18: কোনো একটা matrix-এর JCF তৈরী হয়েছে একটা $J_2(i)$, একটা $J_3(i)$ এবং দুটো $J_3(5)$ নিয়ে। এই matrix-টার eigenvalue-গুলো এবং তাদের algebraic multiplicity-গুলো বার করো। সেই সঙ্গে characteristic polynomial-টাও লেখো। ■

67.1.2 Determinant

যেকোনো triangular matrix-এর বেলাতেই তার diagonal entry-গুলোর গুণফল হয় তার determinant. যেহেতু JCF-ও একটা triangular matrix, তাই সে কথা এখানেও খাটবে।

Example 20: যে matrix-এর JCF তৈরী হয়েছে চারটে $J_2(2+i)$ এবং দুটো $J_1(3)$ দিয়ে, তার determinant কী হবে?

SOLUTION: $\left((2+i)^2\right)^4 \times 3^2 = 9(2+i)^8$. ■

Exercise 19: একটা matrix আছে, যার JCF-এ $J_3(4)$ আছে দুটো, আর $J_5(2i)$ আছে তিনটে। Matrix-টার determinant কত? ■

67.1.3 Geometric multiplicity, nullity, rank

কোন্ eigenvalue-র জন্য কটা Jordan block আছে, সেটাই হল তার geometric multiplicity. এটা আমরা সেই বিন্দুর যুক্তি দিয়ে সহজেই বুঝতে পারি। প্রথম লাইনে বিন্দু থাকে সবচেয়ে বেশী। সুতরাং প্রথম লাইনে বিন্দু যতগুলো, ততগুলোই column হবে। যেমন ধরো এই ছবিতে--

প্রথম লাইনে বিন্দুর সংখ্যা 4, এবং সেটাই হল column-এর সংখ্যা। এদিকে প্রথম লাইনের বিন্দুর সংখ্যা আমরা জানি $\nu(A-\lambda I)$, মানে $\text{gm}(\lambda, A)$.

Example 21: যদি একটা JCF তৈরী হয় একটা $J_1(3)$, দুটো $J_5(4)$ আর তিনটে $J_2(3)$ দিয়ে, তবে এর বিভিন্ন eigenvalue-র geometric multiplicity কী কী?

SOLUTION: এখানে eigenvalue হল 3 আর 4. দেখা যাচ্ছে যে, 3-এর জন্য Jordan block আছে চারটে-- একটা $J_1(3)$ আর তিনটে $J_2(3)$. তাই 3-এর geometric multiplicity হল 4. একইভাবে 4-এর geometric multiplicity হল 2, কারণ ঠিক দুটোই ব্লক আছে 4-এর জন্য।
লক্ষ্য করো, ব্লকগুলোর সাইজ এখানে গুরুত্বপূর্ণ নয়, ওদের সংখ্যাটুকুই খালি কাজে লাগছে। ■

Exercise 20: একটা JCF-এর মধ্যে আছে খালি পাঁচটা $J_3(i)$, দুটো $J_1(i)$, তিনটে $J_4(2)$ এবং দশটা $J_3(-i)$. তবে এর বিভিন্ন eigenvalue-র geometric multiplicity বার করো। ■

কেন কোনো eigenvalue-এর geometric multiplicity তার Jordan block-এর সংখ্যার সমান হবে, সেটা আমরা বিন্দুর গল্প দিয়ে বলেছি। সরাসরি প্রমাণ করাও খুব কঠিন নয়। বোঝার জন্য প্রথমে এমন একটা JCF নেওয়া যাক, যাতে খালি একটাই Jordan block আছে।

Example 22: ধরো $A = J_5(3)$ দিলাম। এর eigenvalue হল 3. এখানে একটাই Jordan block আছে, তাই geometric multiplicity-টা হওয়া উচিত 1. মানে $\nu(A-3I) = 1$. সরাসরি $A-3I$-র nullity বার করে দেখো তো তাই হচ্ছে কিনা।

SOLUTION: এখানে $A-3I$ মানে $J_5(3)$-এর diagonal-এর 3-গুলো বিয়োগ করে শূন্য করে দেওয়া হয়েছে। তাই

$$A - 3I = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

এর প্রথম column-টা স্রেফ null vector, কিন্তু বাকিগুলো linearly independent. সেই জন্য এর rank হবে $5-1=4$. তাই $\nu(A-3I) = 5-4 = 1$. ■

একই যুক্তি খাটবে যেকোনো Jordan block-র ক্ষেত্রেই। তাই λ কোনো eigenvalue হলে তার geometric multiplicity-তে প্রতিটা $J_k(\lambda)$-ই 1 যোগ করবে। এটা ভালো করে ভেবে বুঝে নাও। তাই geometric multiplicity হবে $J_k(\lambda)$-দের মোট সংখ্যা।
একই কায়দায় একটা JCF-এর nullity আর rank-ও বার করা যাবে। nullity বার করার জন্য দেখতে হবে 0 আমাদের matrix-টার একটা eigenvalue কিনা। যদি না হয়, তবে JCF-টা nonsingular, তাই nullity হবে 0. আর যদি 0 একটা eigenvalue হয়, তবে তার geometric multiplicity-ই হবে matrix-টার nullity, কারণ matrix-টার নাম যদি J হয়, তবে 0-র geometric multiplicity মানে হল $\nu(J-0I) = \nu(J)$. তাই nullity আর করার জন্য খালি গুণে দেখতে হবে 0-র জন্য কতগুলো Jordan block আছে।

Example 23: ধরো এই JCF-টা দিলাম, $\text{diag}(J_1(4), J_2(4), J_1(3), J_3(0), J_2(i))$. এর nullity কত? Rank-ই বা কত?

SOLUTION: এখানে একটা eigenvalue হল 0. যেহেতু 0-র জন্য খালি একটাই Jordan block আছে, তাই nullity-ও হল 1. এদিকে column-এর সংখ্যা 9. সুতরাং rank হবে $9-1=8$. ■

Exercise 21: আগের অংকটাই আরেকবার করো, কিন্তু এবার এই JCF-টা নিয়ে--

$$\mathrm{diag}(J_3(5), J_1(2), J_1(5), J_3(4), J_3(2+3i)).$$

■

67.1.4 Minimal polynomial

Minimal polynomial কাকে বলে মনে আছে কি? না থাকলে একবার চট্ করে মনে করিয়ে দিই। কোনো matrix-এর minimal polynomial হল এমন একটা nonzero polynomial, যেটা সেই matrix-এর উপরে লাগালে O পাওয়া যায়। অবশ্য এরকম অসংখ্য polynomial আছে। আমরা তাদের মধ্যে একটা বিশেষ polynomial-কেই খালি minimal polynomial আখ্যা দিই--যেটা এদের মধ্যে সবচেয়ে কম degree-র, এবং যেটা monic, অর্থাৎ সবচেয়ে বড় power-টার coefficient হল 1. সব square matrix-এরই যে ঠিক একটাই minimal polynomial আছে, সে সব আমরা আগেই আলোচনা করেছি। এবার দেখব কোনো matrix-এর JCF দেওয়া থাকলে, তা থেকে চোখে দেখেই কী করে ধাঁ করে তার minimal polynomial বার করে দেওয়া যায়।
তোমাকে খালি দেখে নিতে হবে প্রত্যেকটা eigenvalue-র জন্য সবচেয়ে বড় Jordan block-টার সাইজ কত। ধরো eigenvalue-গুলো হল $\lambda_1 \neq \cdots \neq \lambda_k$ এবং এদের জন্য সবচেয়ে বড় Jordan block-দের সাইজ হল যথাক্রমে $p_1, ..., p_k$. তবে minimal polynomial হবে

$$m(\lambda) = (\lambda - \lambda_1)^{p_1} \cdots (\lambda - \lambda_k)^{p_k}.$$

এটা কেন হয়, সেটা এক্ষুণি বলব। কিন্তু আগে কয়েকটা অংক করে জিনিসটা মাথায় ঢুকিয়ে নিই।

Example 24: যে matrix-এর JCF তৈরী হয় $J_2(2)$, $J_3(4)$, $J_1(2)$ এবং $J_2(4)$ দিয়ে, তার minimal polynomial কী হবে?

SOLUTION: এখানে 2-এর জন্য সবচেয়ে বড় Jordan block-এর সাইজ হল 2. একইভাবে 4-এর জন্য সবচেয়ে বড় Jordan block-এর সাইজ হল 3. তাই minimal polynomial-টা হবে

$$m(\lambda) = (\lambda - 2)^2(\lambda - 4)^3.$$

■

Exercise 22: আগের অংকটাই কর নীচের JCF-টা নিয়ে--

$$\mathrm{diag}(J_2(i), J_1(i), J_1(-i), J_2(-i), J_3(5), J_4(0), J_1(0)).$$

■

এবার বলি এই কায়দাটা কেন কাজ করে। এর পিছনে $J_k(0)$-দের একটা ধর্ম লাগবে। বোঝানোর জন্য ধরো $k = 5$ নিচ্ছি। তবে Jordan block-টা দেখিয়েছি Fig 4-এ। যদি একে square করো, তবে পাবে Fig 5-এর matrix-টা। দেখলে, 1-এর হেলানো সারিটা কেমন এক ঘর উপরে সরে গেল? ফলে 1-এর সংখ্যা কমে গেছে। যদি cube বার করতে, তবে পেতে Fig 6. লক্ষ্য কর যে, 1-এর সারিটা আরও এক ঘর উঠে গেল। অনুমান করতে পারছ নিশ্চয়ই যে, $\big(J_5(0)\big)^4$-এর বেলায় সারিটা আরও এক ঘর উঠবে, ফলে একখানা 1 খালি টিমটিম করে উপরের ডানদিকের ঘরে আটকে থাকবে। যেই $\big(J_5(0)\big)^5$ বার করবে, অমনি সেই সবেধন নীলমণি 1-টাও স্বর্গে উঠে যাবে, ফলে $\big(J_5(0)\big)^5 = O$ হবে।
এই ব্যাপারটা যেকোনো $J_k(0)$-র জন্যই হবে-- সবচেয়ে ছোটো যে n-এর জন্য $\big(J_k(0)\big)^n = O$ হবে, সেটা হল $n = k$. নীচের অংকে সেটাই প্রমাণ করতে বলেছি। যদি প্রমাণটা আপাততঃ সাধে বা সাধ্যে না কুলোয়, তবে জিনিসটা মাথায় রেখে প্রমাণ ছাড়াই বাকি আলোচনাটুকু পড়ো।

$$J_5(0) = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$$

Fig 4

$$\big(J_5(0)\big)^2 = \begin{bmatrix} 0 & 0 & 1 & & \\ & 0 & 0 & 1 & \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}$$

Fig 5

$$\big(J_5(0)\big)^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & \\ & 0 & 0 & 0 & 1 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}$$

Fig 6

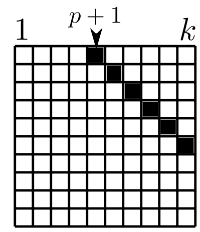


Fig 7

Exercise 23: Prove that the smallest positive n for which $\big(J_k(0)\big)^n = O$ is $n = k$.
HINT: এখানে তোমাকে induction-এর সাহায্যে দেখাতে হবে যে, যেকোনো $p \in \mathbb{N}$-এর জন্যই $\big(J_k(0)\big)^p$-এর মধ্যে 1-এর diagonal সারিটা p-এর সঙ্গে সরে সরে যাবে। যেমন যদি $p \in \{1, ..., k\}$ হয়, তবে $\big(J_k(0)\big)^p$-কে দেখতে হবে Fig 7-এর মত। এখানে কালো ঘরে 1 আছে, সাদা ঘরে 0. অংকের ভাষায়, $\big(J_k(0)\big)^p$-এর (i,j)-th entry হবে--

$$\begin{cases} 1 & \text{if } j = i + p \\ 0 & \text{otherwise} \end{cases}.$$

যেই $p \geq n$ হবে, তখন কোনো (i,j)-র জন্যই আর 1 হবার শর্তটা পালিত হবে না, ফলে পুরো matrix-টাই শূন্য হয়ে যাবে। ■

এবার ধরো একটা JCF আছে, যার মধ্যে খালি দুটো Jordan block, $J_3(5)$ এবং $J_7(5)$, দুজনেই একই eigenvalue-র জন্য। তাহলে $J - 5I = \text{diag}(J_3(0), J_7(0))$ হবে। আমরা জানি যে, একটা block diagonal matrix-এর power নিলে, block-গুলোর ঘাড়ে power-টা চাপে। সুতরাং $(J - 5I)^k = \text{diag}((J_3(0))^k, (J_7(0))^k)$ হবে। যেই $k = 3$ হবে, অমনি $(J_3(0))^k$-টা আউট হয়ে যাবে। কিন্তু $(J_7(0))^k$ অংশটা তখনও ব্যাটিং করে যাচ্ছে। ওটা আউট হবে $k = 7$-এর মাথায়। দুজনেই তখন আউট, ফলে $(J - 5I)^k = O$ হবে প্রথম $k = 7$-এর জন্য। তাই minimal polynomial হবে $(\lambda - 5)^7$. লক্ষ্য করো 7-টা কী করে এল। ওটা হল সবচেয়ে বড় Jordan block-টার সাইজ, যে Jordan block-টা শেষ পর্যন্ত ব্যাটিং করে গেছিল।

67.2 এমন matrix বানাও, যাতে...

এক ধরণের অংক আছে, যেগুলো JCF ছাড়া করা একরকম অসম্ভব বললেই হয়। এই সব অংকে কিছু শর্ত দিয়ে এমন একটা matrix বানাতে বলা হয় যেটা সেই শর্তগুলো পালন করে, অথবা দেখাতে বলা হয় যে, সেরকম matrix-এর অস্তিত্ব নেই। শর্তগুলো যদি rank, nullity, eigenvalue, algebraic multiplicity ইত্যাদি বৈশিষ্ট্য নিয়ে হয়, মানে যেসব বৈশিষ্ট্য similar matrix-দের বেলায় একই হয়, তবে এই সব অংকের অমোঘ সমাধান হল JCF নিয়ে চিন্তা করা। একটা উদাহরণ দেখাই।

Example 25: এমন একটা matrix দিতে পারো, যার একটা eigenvalue হল 5, এবং সেই eigenvalue-র algebraic multiplicity হচ্ছে 10 আর geometric multiplicity হচ্ছে 3?
SOLUTION: অংকটা দেখামাত্র তুমি ব্যাপারটাকে JCF দিয়ে চিন্তা করো--এমন একটা JCF চাইছে, যার 5-এর Jordan block-গুলো আছে তিনটে, আর ওদের মোট সাইজ হল 10. মানে 10-কে যেকোনো 3-টে ভাগে ভেঙে ফেললেই হল, যেমন $10 = 5 + 3 + 2$. তাহলে $\text{diag}(J_5(5), J_3(5), J_2(5))$ হবে একটা উত্তর। ■

এই অংকটা করার পর নীচের অংকটা খুব শক্ত লাগা উচিত নয়।

Exercise 24: আমরা জানি যে, যেকোনো eigenvalue-র জন্যই

$$1 \leq \text{geometric multiplicity} \leq \text{algebraic multiplicity}.$$

যদি তোমাকে যা খুশী কিছু $a_1, ..., a_k \in \mathbb{N}$ এবং যা খুশি $g_1, ..., g_k \in \mathbb{N}$ দিই, যেখানে $\forall i \quad 1 \le g_i \le a_i$, তবে কি এমন matrix থাকতে বাধ্য, যার ঠিক k-খানা আলাদা আলাদা eigenvalue আছে, যাদের algebraic multiplicity-গুলো হল a_i, আর geometric multiplicity-গুলো হল g_i? ■

Exercise 25: এমন একটা matrix বানাও, যার একটা eigenvalue হল 3, যার minimal polynomial-টা $(\lambda-3)^5$ দিয়ে ভাগ যায় না, এবং 3-এর algebraic multiplicity আর geometric multiplicity হল যথাক্রমে 9 আর 2.

HINT:

JCF দিয়ে ভাবলে শর্তগুলো হল--diagonal-এ মোট নয়টা 3 আছে, সবচেয়ে বড় 3-এর ব্লকের সাইজ হতে পারে $5-1=4$, আর ব্লকের সংখ্যা 2. এবার চিন্তা করো। ■

একই কায়দায় নীচের অংকটাও হবে।

Exercise 26: Let λ_0 be an eigenvalue with algebraic and geometric multiplicities a and g, respectively. Let $k = \lceil \frac{a}{g} \rceil$, the smallest integer $\ge \frac{a}{g}$. Then show that $(\lambda - \lambda_0)^k$ must divide the minimal polynomial. ■

এই অংকটা হয়ে গেলে এবারের অংকটাও খুব শক্ত লাগা উচিত নয়, যদিও লেখাটা একটু প্যাঁচের।

Exercise 27: Let λ_0 be an eigenvalue of a complex matrix, and its algebraic multiplicity and geometric multiplicity be, respectively, a and g. Also let $(\lambda-\lambda_0)^k$ be the highest power of $(\lambda-\lambda_0)$ in its minimal polynomial $m(\lambda)$. Then show that k can take any value in $\left\{\lceil \frac{a}{g} \rceil, ..., a-g+1\right\}$, but no value outside this set.

HINT:

আগের অংক থেকেই $k \ge \lceil \frac{a}{g} \rceil$ জানা আছে। প্রথমে দেখো $k > a-g+1$ হতে পারে কিনা। সেটা যে পারে না, সেটা দেখানো কঠিন নয়। এরপর দেখাতে হবে যেকোনো $k \in \left\{\lceil \frac{a}{g} \rceil, ..., a-g+1\right\}$-এর জন্যই এই কাজটা করা যাবে--

> মোট a-খানা λ_0-কে g-টা ব্লকে ভাগ করতে হবে। সবচেয়ে বড় ব্লকের সাইজ রাখতে হবে k.

প্রথমেই k-খানা সরিয়ে রাখো, সেটা হবে আমাদের প্রথম ব্লক। যদি $g=1$ হয়, তবে গল্প এখানেই শেষ। যদি $g \ge 2$ হয়, তবে এবার বাকি $a-k$-খানা λ_0-কে যথাসম্ভব সমান সমান ভাগ করে বাকি $g-1$-টা ব্লক বানাও। এই কাজটা গুছিয়ে লেখা খানিক শক্ত। কীভাবে আমরা eigenvalue-গুলো কয়েকটা গুচ্ছে ভাঙছি, সেটা গ্রাফের মত ছবি এঁকে নিলে সুবিধা হবে (Fig 8)। এখানে x-axis বরাবর কটা গুচ্ছ গুনছি, আর y-axis বরাবর হিসেব রাখছি এ যাবত যে কয়টা গুচ্ছ হয়েছে তাদের মধ্যে মোট কয়টা eigenvalue আছে। যেমন Fig 8-এ চারটে বিন্দু আছে--

$$(1,2), (2,3), (3,5), (4,6).$$

এর মানে মোট চারটে গুচ্ছ আছে, প্রথম গুচ্ছে দুটো, প্রথম দুটো মিলিয়ে মোট তিনটে , প্রথম তিনটে মিলিয়ে মোট পাঁচটা, এবং সবগুলো গুচ্ছ মিলিয়ে মোট ছয়টা eigenvalue আছে। শেষের বিন্দুটা থেকে এটাও বোঝা যাচ্ছে যে, eigenvalue-টার algebraic multiplicity হল 6 এবং geometric multiplicity হল 4. এবার Fig 9-এর দিকে তাকাও। আমরা প্রথম গুচ্ছে k-খানা eigenvalue নিয়ে P বিন্দুতে এসেছি। আমাদের যেতে হবে Q বিন্দুতে। সোজা গেলে যে পথে যেতাম, সেটা লাইন টেনে দেখিয়েছি। সমস্যা হল এটা সব সময়ে integer বিন্দু দিয়ে যায় নি। কিন্তু একটা গুচ্ছে তো আর ভগ্নাংশ-সংখ্যক eigenvalue নিতে পারি না। তাই আমরা প্রতিক্ষেত্রে ঠিক নীচের integer-টা নিয়েছি। মানে i-th গুচ্ছে $[(i-1)s]-[(i-2)s]$-খানা eigenvalue রাখছি, যেখানে $s = \frac{a-k}{g-1}$ হল লাইনটার slope, এবং $[x]$ হল largest integer $\le x$. এখানে $i = 2, ..., g$. মাথা গুলিয়ে যাচ্ছে? ধীরে ধীরে চিন্তা করো। তোমার কাজ এখন খালি এইটা দেখানো

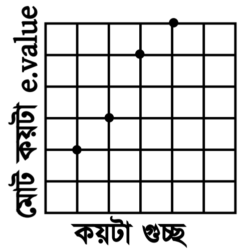


Fig 8

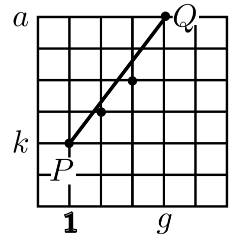


Fig 9

যে, এইভাবে এগোলে প্রতিটা গুচ্ছের সাইজই 1 থেকে k-র মধ্যে থাকবে। ■

Example 26: Find all possible Jordan canonical forms for the matrix whose characteristic polynomial is $(t-2)^4(t-5)^3$ and minimal polynomial is $(t-2)^2(t-5)^3$.[2] **(2021.CC12.3cii)**

SOLUTION:

> Here the algebraic multiplicity of 2 is 4. So 2 appears exactly 4 times in the diagonal.
>
> Also since $(t-2)$ has power 2 in the minimal polynomial, the maximum size of Jordan block with diagonals 2, is 2.

তার মানে এই eigenvalue-টার জন্য আমাদের কাজ হল এমনভাবে Jordan block বানানো, যাতে ওদের মোট সাইজ হয় 4, আর সবচেয়ে বড় সাইজটা হয় 2. কাজটা খালি দুভাবেই করা যায়, $4 = 2 + 2$ এবং $4 = 2 + 1 + 1$.

> So we have either $J_2(2), J_2(2)$ or $J_2(2), J_1(2), J_1(2)$.

এবার একই কাজ করতে হবে অন্য eigenvalue-টা নিয়ে--

> Again, the eigenvalue 5 occurs with algebraic multiplicity 3.
>
> Since we have $(t-5)^3$ in the minimal polynomial, the largest size is 3.
>
> So there must be a Jordan block $J_3(5)$.

এবার কতগুলো JCF সম্ভব, সেটা বার করতে হবে।

> So the total number of JCF not taking into account ordering of the blocks is 2.
>
> However, if we take the order of blocks into account, then the answer is $\frac{3!}{2!1!} + \frac{4!}{1!2!1!} = 15$.

এইটা কী করে হল, বুঝলে কী? মনে করো, তোমাকে দুটো A আর একটা B-কে পরপর সাজাতে হবে। কতভাবে করা যাবে? উত্তর হল $\frac{3!}{2!1!}$. এটা হায়ার সেকণ্ডারীতেই আমরা শিখে থাকি। এখানে A-র জায়গায় $J_2(4)$ আর B-র জায়গায় $J_3(5)$ কল্পনা করো। এইটা হল প্রথম কেসটা। দ্বিতীয় কেসে $J_2(2), J_1(2), J_1(2)$ আর একটা $J_3(5)$ আছে, মানে চারটে জিনিস,

যাদের মধ্যে দুজন একইরকম। সেখান থেকেই $\frac{4!}{1!2!1!}$ এসেছে। ■

Exercise 28: Find all possible Jordan canonical forms for the complex matrix whose characteristic polynomial is $(t-1)^4(t^2+1)^2(t-2)^3$ and minimal polynomial is $(t-1)(t-2)^2(t^2+1)$. ■

67.3 Real field-এর উপরে JCF

JCF বার করার সময়ে আমরা ধরে নিই যে, কাজটা হচ্ছে $\mathbb{C}$-এর উপরে। বস্তুতঃ eigenvalue নিয়ে কাজ করার অংকে সেটাই দস্তুর। কারণ $\mathbb{C}$-এর উপরে যেকোনো polynomial-এরই root থাকে, তাই যেকোনো $n \times n$ সাইজের matrix-এরই ঠিক n-খানা eigenvalue পাওয়া যায় (algebraic multiplicity-দের হিসেবে নিয়ে)। কিন্তু যদি আমাদের matrix-টার প্রতিটা entry-ই real হয়, তবে কি খালি real সংখ্যা দিয়ে তৈরী কোনো JCF পাওয়া সম্ভব? সব সময়ে যে সম্ভব নয়, সে তো বুঝতেই পারছ, কারণ এমন real matrix আছে, যার অন্ততঃ একটা non-real eigenvalue আছে[3], যেমন $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$-র eigenvalue হল $\pm i$. অতএব যাই JCF নাও না কেন, তার diagonal-এ তো i আর $-i$ না লিখে পথ নেই!

যদিও real সংখ্যা দিয়ে তৈরী JCF পাওয়া নাও যেতে পারে, তাও খালি real সংখ্যাদের দিয়েই একটা জিনিস বানানো সম্ভব, যেটা ঠিক JCF না হলেও কতকটা ওইরকমই দেখতে, এবং JCF-কে সামান্য এদিক-ওদিক করেই তৈরী। একে অনেক সময়ে লোকে **JCF over** $\mathbb{R}$ বলে থাকে। এর জন্য Jordan block-এর মত আরেকধরণের ব্লকের প্রয়োজন হবে, যাকে **Jordan block over** $\mathbb{R}$ বলা যেতে পারে। ধরো $z = \lambda + \mu i \in \mathbb{C}$, যেখানে $\lambda, \mu \in \mathbb{R}$. তবে z-এর জন্য একটা Jordan block over $\mathbb{R}$ হল এইরকম একটা matrix--

$$R_k(z) = \begin{bmatrix} T & I & & & \\ & T & I & & \\ & & \ddots & \ddots & \\ & & & T & I \\ & & & & T \end{bmatrix},$$

যেখানে $T = \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}$, এবং এখানে k-খানা diagonal block আছে। লক্ষ্য করো যে, $R_k(z)$-এর সাইজ কিন্তু $2k \times 2k$.

Example 27: $R_2(2+3i)$ লেখো।

SOLUTION:

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ -3 & 2 & 0 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & -3 & 2 \end{bmatrix}.$$

■

Exercise 29: $R_1(i)$ লেখো। এবার $R_1(-i)$ লেখো। ■

এবার JCF over $\mathbb{R}$-এর ব্যাপারটা একটা theorem-এর আকারে লেখা যাক।

[3]সাবধান, complex বললাম না, বললাম non-real, কারণ real সংখ্যাদেরও complex বলে ভাবা যায়।

JCF over $\mathbb{R}$

Any real square matrix is similar to a real, square, block diagonal matrix where each diagonal block is of the form $J_k(\lambda)$ or $R_k(z)$ for some $\lambda \in \mathbb{R}$, $z \in \mathbb{C}$ and $k \in \mathbb{N}$. For any real square matrix this block diagonal matrix is unique up to the order of these blocks, except that an $R_k(z)$ may be replaced by $R_k(\overline{z})$.

JCF over $\mathbb{R}$ বার করা সাধারণ JCF বার করার মতই। প্রথমে field-টাকে $\mathbb{C}$ কল্পনা করে সব eigenvalue এবং তাদের algebraic multiplicity বার করে নিতে হয়। যেসব eigenvalue দেখবে real, তাদের জন্য আগে যা Jordan block পেতে, এখানেও সেগুলোই থাকবে। যদি z কোনো non-real eigenvalue হয়, তবে $\overline{z}$-ও আরেকটা eigenvalue হবে, এবং ওদের Jordan block-রা একইরকম হবে, মানে সংখ্যায় এবং সাইজে। এমনটা কেন হবে, সেটা একটু ভেবে নিতে পারো। যদি স্পষ্ট না হয়, চিন্তা নেই। এই অধ্যায়ের শেষে প্রমাণ করার সময়ে বুঝে যাবে।
যখনই $J_k(z)$ থাকবে, অমনি তার সাথে তাল রেখে একটা $J_k(\overline{z})$-ও থাকবে। এই দুজনকেই সরিয়ে দিয়ে তুমি তার জায়গায় একখানা $R_k(z)$ (বা $R_k(\overline{z})$) লিখে দেবে, ব্যস! লক্ষ করো, $J_k(z)$ আর $J_k(\overline{z})$ দুজনে diagonal block হিসেবে মোট $2k$-খানা row এবং column দখল করেছিল। $R_k(z)$ বা $R_k(\overline{z})$ একাই সেই পরিমাণ জায়গা দখল করবে। একটা উদাহরণ দেখলে বুঝতে সুবিধা হবে।

Example 28: ধরো কোনো real matrix-এর JCF হল

$$\mathrm{diag}\big(J_2(1), J_3(1), J_2(2+3i), J_1(2+3i), J_2(2-3i), J_1(2-3i)\big).$$

এর জন্য JCF over $\mathbb{R}$ বার করো।
SOLUTION: এখানে Jordan block-গুলো হল $J_2(1)$, $J_3(1)$, $J_2(2+3i)$, $J_1(2+3i)$, $J_2(2-3i)$ এবং $J_1(2-3i)$. এদের মধ্যে প্রথম দুটো অপরিবর্তিত থাকবে। $J_2(2+3i)$ আর $J_2(2-3i)$-এর বদলে নেব $R_2(2+3i)$. একইভাবে $J_1(2+3i)$ এবং $J_1(2-3i)$ মিলিয়ে নেব $R_1(2+3i)$. সুতরাং একটা real JCF হবে এইরকম--

$$\mathrm{diag}\big(J_2(1), J_3(1), R_2(2+3i), R_1(2+3i)\big).$$

এখানে $R_2(2-3i)$ আর $R_1(2-3i)$-ও নেওয়া যেত। ■

লক্ষ্য করো, যদি কোনো real matrix-এর সবগুলো eigenvalue-ই real হয়, তবে এমনি JCF-ই তার JCF over $\mathbb{R}$-ও হবে। নীচের অংকে সেটাই তোমার কাজে লাগবে।

Exercise 30: Find the Jordan normal form of $\begin{bmatrix} 4 & -1 & 1 \\ 4 & 0 & 2 \\ 2 & -1 & 3 \end{bmatrix}$ over the field of real.[5]
(2022.CC12.3f)
HINT: এই অংকের Jordan normal form বলতে JCF বোঝাচ্ছে। Matrix-টার eigenvalue-গুলো সব বার করলেই দেখবে সকলেই real. তাই সাধারণ JCF-এই কাজ চলে যাবে। আর হ্যাঁ, এখানে বলা উচিত ছিল 'a' Jordan normal form, যদিও এখানে 'the' Jordan normal form বলেছে। ■

DAY 68
Rational canonical form (RCF)

JCF বার করার জন্য আমরা $\mathbb{C}$-এর উপরে কাজ করেছিলাম। তার কারণ $\mathbb{C}$ এমন একটা field, যার উপর যেকোনো polynomial-কেই সম্পূর্ণ factorise করা যায়। সেই জন্য eigenvalue বার করার কোনো সমস্যা ছিল না। কিন্তু সব field-এর ক্ষেত্রে সেটা হয় না, যেমন $\mathbb{R}$-এর ক্ষেত্রেই হয় না, বা $\mathbb{Q}$-এর ক্ষেত্রেও হয় না। স্বাভাবিকভাবেই প্রশ্ন ওঠে যে, JCF-এর এমন কোনো সংস্করণ আছে কিনা, যেটা যেকোনো field-এর উপরেই কাজ করে। উত্তর হল--হ্যাঁ আছে। Rational canonical form হল এরকমই একটা জিনিস। JCF-এর বেলায় যেমন Jordan block বলে একটা বিশেষ ধরণের matrix-এর কথা শিখতে হয়েছিল, rational canonical form-এর বেলাতেও তেমনি একধরণের matrix-এর কথা শিখতে হবে, যাকে বলে **companion matrix**.

68.1 Companion matrix

68.1.1 সংজ্ঞা

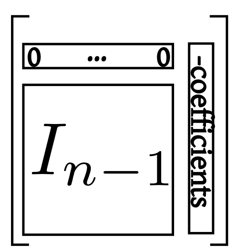
DEFINITION: **Companion matrix**

Let $b(x) = b_0 + b_1 x + \cdots + b_{k-1}x^{k-1} + x^k$ be any monic polynomial over a field, F. Then its **companion matrix** is the matrix

$$\begin{bmatrix} 0 & 0 & \cdots & \cdots & \cdots & -b_0 \\ 1 & 0 & \cdots & \cdots & \cdots & -b_1 \\ 0 & 1 & \cdots & \cdots & \cdots & -b_2 \\ 0 & 0 & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & -b_{k-1} \end{bmatrix}.$$

লক্ষ্য করো, এখানে field-টা নিয়েছি F, যেটা $\mathbb{R}$ বা $\mathbb{C}$ হতেও পারে, নাও পারে। বস্তুতঃ আজকে আমাদের পুরো আলোচনাটাই হবে general কোনো field-এর উপরে। যখন আমরা বলছি "a polynomial over a field F", তার মানে polynomial-টার সবগুলো coefficient-ই F-এর সদস্য।
Companion matrix-দের গঠনটা দেখিয়েছি Fig 10-এ। নীচের বাঁদিকে একটা $(n-1)\times(n-1)$ সাইজের identity matrix, তার উপরে একটা শূন্যর ঢাকনা চাপানো। ডানদিকের column-টা জুড়ে রয়েছে polynomial-এর coefficient-গুলোর negative-রা। উপর থেকে যত নীচে যাবে, ততই power-গুলো বাড়বে। সবচেয়ে উপরে constant term-টার negative, তার নীচে linear term-টার coefficient-এর negative, ইত্যাদি। এইভাবে চলতে চলতে একেবারে নীচে

Fig 10



থাকবে x^{n-1}-এর coefficient-এর negative. লক্ষ্য করো x^n-এর coefficient হল 1 (কারণ monic polynomial), কিন্তু এই coefficient-টা companion matrix-এ কোথাও ব্যবহৃত হয়নি। আরো লক্ষ্য করো যে, matrix-টার সাইজ হল $\deg(b(x))$-এর সমান।
একটা উদাহরণ করে গঠনটা হজম করা যাক।

Example 29: $x^2 - 3x + 5$-এর companion matrix কী হবে?

SOLUTION: আমাদের polynomial-টার degree হল 2. তাই matrix-টার সাইজ হবে 2×2. Matrix-টা হবে

$$\begin{bmatrix} 0 & -5 \\ 1 & 3 \end{bmatrix}.$$

■

Exercise 31: $x^3 + x + 1$-এর companion matrix বানাও। ■

Exercise 32: $x + a$-র companion matrix কী হবে? এখানে a হল যা খুশি একটা scalar. ■

হঠাৎ খামোখা এরকম একটা matrix বানানোর কারণ কী? আসলে এই matrix-টার সঙ্গে ওই polynomial-টার ওতপ্রোত সম্পর্ক আছে। এবার সেটাই আলোচনা করব।

68.1.2 Characteristic এবং minimal polynomial

কোনো monic polynomial-এর companion matrix-এর characteristic polynomial হবে ওই polynomial-টাই। এটা আমরা একটু পরেই প্রমাণ করব, আগে কিছু উদাহরণ নিয়ে দেখি।

Example 30: এক্ষুণি একটা অংকে আমরা দেখলাম যে, x^2-3x+5-এর companion matrix হল $C = \begin{bmatrix} 0 & -5 \\ 1 & 3 \end{bmatrix}$.

দেখাও যে, এই matrix-টার characteristic polynomial হবে ওই polynomial-টাই।
SOLUTION: Characteristic polynomial-টা হবে

$$\det(xI - C) = \begin{vmatrix} x-0 & 5 \\ -1 & x-3 \end{vmatrix} = x(x-3) + 5 = x^2 - 3x + 5,$$

ঠিক যেমনটা দরকার ছিল। ■

Exercise 33: এই matrix-টার characteristic polynomial বার করো--

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

■

এবার আমরা ব্যাপরটা গুছিয়ে লিখে প্রমাণ করব।

THEOREM

Let $p(x)$ be any monic polynomial. Let C be its companion matrix. Then the characteristic polynomial of C must be $p(x)$, *i.e.*,

$$\det(xI - C) = p(x).$$

Proof:

> We shall use induction on $\deg(p(x))$.
> Basis: If $\deg(p(x)) = 1$, then $p(x) = x + a$, for some scalar a.
> So $C = [-a]$, with characteristic polynomial $\det(xI - C) = \det[x + a] = x + a = p(x)$, as required.

এবার induction hypothesis-টা লিখে নিই--

> Hypothesis: We assume the result for $\deg(p(x)) = 1, 2, ..., n$ for some $n \in \mathbb{N}$.

অবশেষে induction step. এখানেই আসল পরিশ্রমটা হবে--

> Step: Shall prove for $\deg(p(x)) = n + 1$.
> Let $p(x) = a_0 + a_1 x + \cdots + a_n x^n + x^{n+1}$. Then
>
> $$\det(xI - C) = \begin{vmatrix} x & & & & a_0 \\ -1 & x & & & a_1 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & x & a_{n-1} \\ & & & -1 & a_n \end{vmatrix}.$$

ফাঁকা ঘরগুলোতে শূন্য আছে। আমরা এবার এর উপরে Laplace expansion লাগাব, প্রথম row বরাবর। যেহেতু প্রথম row-তে দুজন ছাড়া সকলেই শূন্য, তাই expansion-টায় খালি দুটো term-ই থাকবে--

> Expanding by the first row, this is
>
> $$\det(xI - C) = x \begin{vmatrix} x & & & & a_1 \\ -1 & x & & & a_2 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & x & a_{n-1} \\ & & & -1 & a_n \end{vmatrix} + (-1)^{1+(n+1)} a_0 \begin{vmatrix} -1 & x & & \\ & \ddots & \ddots & \\ & & \ddots & x \\ & & & -1 \end{vmatrix}.$$

এইখানটায় হয়তো খাবি খেতে পারো। ভালো করে বুঝে নিয়ে এগোও। ওই $(-1)^{1+(n+1}$-টা কেন এসেছে? কারণ আমাদের matrix-টা $(n+1) \times (n+1)$ সাইজের, এবং a_0 রয়েছে $(1, n+1)$ অবস্থানে।

প্রথম determinant-টার গঠনটা বোঝানোর জন্য অনেক ডট্ ডট্ দিতে হয়েছে, তাই ওটার সাইজ একটু বড় দেখাচ্ছে। আসলে কিন্তু দুটো determinant-ই $n \times n$ সাইজের matrix-এর। এদের মধ্যে প্রথমটা ফের একটা companion matrix জাতীয় জিনিস, তাই ওটার উপরে induction hypothesis-টা খাটানো যাবে।

> The first determinant is $\det(xI - D)$, where
>
> $$D = \begin{bmatrix} 0 & & & & -a_1 \\ 1 & x & & & -a_2 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & 0 & -a_{n-1} \\ & & & 1 & -a_n \end{bmatrix}$$
>
> is the companion matrix of $a_1 + a_2x + \cdots + a_nx^{n-1} + x^n$.
> Hence, by the induction hypothesis, the first determinant is $a_1 + a_2x + \cdots + a_nx^{n-1} + x^n$.

দ্বিতীয় determinant-টা সহজ জিনিস, কারণ ওটা একটা triangular matrix-এর determinant. তাই স্রেফ diagonal entry-গুলোকে গুণ করে গেলেই হবে। সবগুলো diagonal entry-ই -1, এবং এরকম n-খানা আছে। তাই--

> Also, the second determinant is $(-1)^n$.

সুতরাং দুয়ে মিলে হল--

> Hence the characteristic polynomial of C is
>
> $$x(a_1 + a_2x + \cdots + a_nx^{n-1} + x^n) + (-1)^{1+(n-1)+n}a_0 = a_0 + a_1x + \cdots + a_nx^n + x^{n+1},$$
>
> as required.

[Q.E.D]

Companion matrix-এর আরেকটা গুণও আছে। সেটা হল $p(x)$-এর companion matrix-এর minimal polynomial-টাও হবে সেই $p(x)$. এই জিনিসটা আমাদের শীঘ্রই কাজে লাগবে।

THEOREM

Let $p(x)$ be any monic polynomial. Let C be its companion matrix. Then the minimal polynomial of C must be $p(x)$.

Proof: প্রমাণটা ছোটোই, কিন্তু একটু ঘুরপথে। Minimal polynomial হল সেই annihilating polynomial, যেটা monic এবং যার degree সবচেয়ে কম। Cayley-Hamilton-এর সুবাদে আমরা জানি যে, characteristic polynomial এদের মধ্যে প্রথম দুটো শর্ত এমনিতেই পালন করে। এর degree হল matrix-টার সাইজ। সুতরাং যদি দেখাতে পারি যে, matrix-টার সাইজের চেয়ে ছোটো degree-র কোনো non-zero annihilating polynomial নেই, তবেই প্রমাণটা হয়ে যাবে।

> Let $n = \deg(p(x))$.
> Let $\{\vec{e}_1, ..., \vec{e}_n\}$ be the canonical basis of F^n, where F is the underlying field.

লক্ষ্য করো, $\vec{e}_i$-গুলো সবই 1 আর 0 দিয়ে তৈরী। যেহেতু যেকোনো field-এর মধ্যেই 0 আর 1 থাকে, তাই canonical basis-এর ধারণাটা যেকোনো F^n-এর জন্যই খাটে।

> Then for any matrix, A, we have $A\vec{e}_i = i$-th column of A.

এই তথ্যটা linear algebra-র অনেক জায়গাতেই খুব কাজে আসে।

> Now, the first $n-1$ columns of C are $\vec{e}_2, ..., \vec{e}_n$.
> Hence for any matrix, A, the first $n-1$ columns of AC are the last $n-1$ columns of A.

এই জায়গাটা কি বোঝা গেল? না বোঝা গেলে C-কে এইভাবে ভাবো, $C = \begin{bmatrix} \vec{e}_2 & \cdots & \vec{e}_n & \vec{*} \end{bmatrix}$. শেষের column-টাকে $\vec{*}$ লিখেছি, মানে আপাততঃ ওটাকে নিয়ে মাথা ঘামাচ্ছি না। আমরা প্রথম অধ্যায়ে block matrix-দের গুণ শিখেছিলাম। সেই অনুযায়ী $AC = \begin{bmatrix} A\vec{e}_2 & \cdots & A\vec{e}_n & A\vec{*} \end{bmatrix}$ হবে। সেখান থেকেই ব্যাপারটা আসছে।

> Using this repeatedly, we see that the first column of C^k is $\vec{e}_{k+1}$ for $k = 1, ..., n-1$.

একটা উদাহরণ নিয়ে দেখো।

$$C = \begin{bmatrix} 0 & 0 & 0 & * \\ 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{bmatrix}, \quad C^2 = \begin{bmatrix} 0 & 0 & * & * \\ 0 & 0 & * & * \\ 1 & 0 & * & * \\ 0 & 1 & * & * \end{bmatrix}, \quad C^3 = \begin{bmatrix} 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 1 & * & * & * \end{bmatrix}.$$

লক্ষ্য করো কীভাবে $*$-গুলো সাম্রাজ্য বিস্তার করছে, এবং শুরুর $\vec{e}_i$-গুলো একঘর একঘর করে সরছে।

> Also the first column of $C^0 = I$ is $\vec{e}_1$.
> Hence for any polynomial $q(x) = b_0 + b_1 x + \cdots + b_{n-1}x^{n-1}$ with $\deg(q(x)) \leq n-1$, the first column of $q(C)$ must be
>
> $$\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

কারণ $q(C) = b_0 I + b_1 C + \cdots + b_n C^n$-এর প্রথম column হবে

$$\sum_0^{n-1} b_k (C^k\text{-এর প্রথম column}) = \sum_0^{n-1} b_k \vec{e}_{k+1}.$$

> Thus, if $q(C) = O$, then we must have $q(x) \equiv 0$.

তার মানে matrix-টার সাইজের চেয়ে ছোটো degree-র কোনো non-zero annihilating polynomial পাওয়া সম্ভব নয়। তাই characteristic polynomial-টাই minimal polynomial হতে বাধ্য। এই জায়গার যুক্তিটা একটু গুছিয়ে লিখতে হবে।

> Hence C has no annihilating polynomial of degree $< n$.
>
> Now, by Cayley-Hamilton theorem, the characteristic polynomial, $p(x)$, is a monic annihilating polynomial of degree n.
>
> So the minimal polynomial, $m(x)$, must also be a monic, polynomial of degree n.

এতেই কিন্তু প্রমাণ হয়ে যাচ্ছে না যে, characteristic polynomial-টাই minimal polynomial হবে। তার জন্য আরেকটু প্যাঁচ আছে--

> We know that the minimal polynomial divides all annihilating polynomials. So $m(x)|p(x)$.
> Since $\deg(m(x)) = \deg(p(x))$, hence $p(x) = c\,m(x)$ for some scalar c.
> Since both $p(x)$ and $m(x)$ are monic, hence $c = 1$.
> So the characteristic polynomial must be the minimal polynomial, as required.

[Q.E.D]

Companion matrix-এর কথা অনেক হল। এবার দেখি rational canonical form তৈরীতে এদের ভূমিকা কী।

68.1.3 Companion matrix থেকে RCF

ঠিক যেমন JCF হল Jordan block দিয়ে তৈরী diagonal matrix, তেমনি rational canonical form হল companion matrix দিয়ে তৈরী diagonal matrix.

DEFINITION: **Rational canonical form**

Let $p_1(x), ..., p_k(x)$ be monic polynomials with positive degrees over a field F such that

$$p_1(x)|p_2(x)| \cdots |p_k(x).$$

Let $C_{p_i(x)}$ be the companion matrix for $p_i(x)$. Then the block diagonal matrix

$$\text{diag}(C_{p_1(x)}, ..., C_{p_k(x)}),$$

is said to be in **rational canonical form**.

সংজ্ঞাটা একটু খটোমটো লাগতে পারে। একটা উদাহরণ দেখলে সুবিধা হবে।

Example 31: ধরো $p_1(x) = 2 + x$ এবং $p_2(x) = 2x + x^2$. এদের দিয়ে তৈরী RCF-টা লেখো।

SOLUTION: এখানে $p_1(x)|p_2(x)$. সুতরাং RCF বানানো যাবে। এখানে

$$C_{p_1(x)} = \begin{bmatrix} 2 \end{bmatrix}, C_{p_2(x)} = \begin{bmatrix} 0 & 0 \\ 1 & -2 \end{bmatrix}.$$

সুতরাং RCF-টা হল

$$\begin{bmatrix} 2 & & \\ & 0 & 0 \\ & 1 & -2 \end{bmatrix}.$$

■

Exercise 34: Find the RCF corresponding to $p_1(x) = x^2 + 1$ and $p_2(x) = (x^2 + 1)(x - 1)$. ■

Rational canonical form decomposition

Every square matrix is similar to a unique rational canonical form.

এই theorem-টা আমরা এ বইতে প্রমাণ করব না। Theorem-টাকে এইভাবেও ভাবা যায়--দুটো matrix পরস্পরের সঙ্গে similar হবে if and only if ওদের RCF একই হয়।
যেহেতু যেকোনো square matrix-ই ঠিক একটাই $\mathrm{diag}(C_{p_1(x)}, ..., C_{p_k(x)})$-র সঙ্গে similar, তাই বুঝতেই পারছ যে, $p_i(x)$-গুলো matrix-টার নিজস্ব ধর্ম। এদের একটা নাম আছে--

DEFINITION: **Invariant factor**

If a square matrix is similar to the rational canonical form

$$\mathrm{diag}(C_{p_1(x)}, ..., C_{p_k(x)}),$$

then the polynomials $p_1(x), ..., p_k(x)$ are called its **invariant factors**.

সুতরাং বুঝতেই পারছ যে, কোনো matrix-এর rational canonical form বার করা মানে তার invariant factor-গুলো বার করা। সেগুলো বার করার কোনো সহজ general পদ্ধতি আছে বলে জানি না। কিন্তু ওদের দুটো ধর্ম বলব, যাদের ব্যবহার করে ছোটোখাটো matrix-এর বেলায় সহজেই rational canonical form বার করে দেওয়া যায়।

68.2 RCF-এর minimal এবং characteristic polynomial

THEOREM

The product of the invariant factors of a matrix equals its characteristic polynomial.

Proof: এটার প্রমাণ অতি সংক্ষিপ্ত। খালি এটুকু লক্ষ্য করলেই হবে যে, কোনো block diagonal matrix-এর characteristic polynomial হল তার block-গুলোর characteristic polynomial-দের গুণফল। এমনটা হয়, কারণ determinant-গুলো গুণ হয়ে যায়, আর characteristic polynomial তো আসলে একটা determinant-ই! **[Q.E.D]**

Exercise 35: যদি $C = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$ হয়, তবে $\text{diag}(C, C)$-র characteristic polynomial কী হবে? ■

THEOREM

The highest degree invariant factor of a matrix must equal its minimal polynomial.

Proof: আমরা আগেই দেখেছি যে, companion matrix-দের minimal polynomial আর characteristic polynomial একই হয়। সেটা থেকেই এই theorem-টা প্রমাণ হবে--

> Let $R = \text{diag}(C_1, ..., C_k)$ be an RCF, where C_i is the companion matrix of a monic polynomial $p_i(x)$ of positive degree, with $p_1(x)|p_2(x)| \cdots |p_k(x)$.

এর মধ্যে সবচেয়ে বেশী degree দেখাই যাচ্ছে $p_k(x)$-এর। তার মানে আমাদের দেখাতে হবে--

> Shall show that the minimal polynomial of R is $p_k(x)$.

এর জন্য R-এর annihilating polynomial-দের সঙ্গে C_i-দের annihilating polynomial-দের একটা সম্পর্ক বার করে রাখলে সুবিধা হবে।

> Let $q(x)$ be any polynomial. Then
>
> $$q(R) = \text{diag}(q(C_1), ..., q(C_k)).$$
>
> Hence $(q(R) = O \iff \forall i \quad q(C_i) = O)$.
>
> So $q(x)$ is an annihilating polynomial for R if and only if it is an annihilating polynomial for all its blocks.

এবার প্রমাণটা অতি সহজেই হবে--

> Now $p_i(x)$ annihilates C_i. Since $p_1(x)|p_2(x)| \cdots |p_k(x)$, hence $p_k(x)$ annihilates all the C_i's.
>
> So $p_k(x)$ annihilates R.
>
> Also no polynomial of less degree can annihilate R, since $p_k(x)$ is the minimal polynomial for C_k.
>
> Hence $p_k(x)$ must be the minimal polynomial for R, as required.

[Q.E.D]

68.3 RCF বার করা

এবার কিছু উদাহরণ দেখাই, যেখানে এই দুটো theorem কাজে লাগিয়ে আমরা কোনো matrix-এর rational canonical form বার করতে পারব। এখানে আমরা ধরে নেব যে, আমরা matrix-টার characteristic polynomial আর minimal polynomial দুটোই জানি।

Example 32: ধরো কোনো matrix-এর characteristic polynomial হল $p(x)$ আর minimal polynomial-ও সেটাই। তবে সেই matrix-এর rational canonical form কী হবে?

SOLUTION:এক্ষেত্রে সবচেয়ে বড় invariant factor-টা হবে এই $p(x)$-টাই। আবার সবগুলো invariant factor-এর গুণফল হতে হবে সেই $p(x)$-ই। যেহেতু invariant factor-দের degree সর্বদা positive হয়, তাই ওই $p(x)$-টাই ওর একমাত্র invariant factor হতে বাধ্য। তার মানে rational canonical form-টায় খালি একটাই block থাকবে, যেটা হবে $p(x)$-এর companion matrix-টা মানে $C_{p(x)}$. ■

এই কায়দাটা খাটবে যখনই characteristic polynomial আর minimal polynomial সমান হবে। এর একটা বিশেষ ক্ষেত্র হল যখন characteristic polynomial-র মধ্যে প্রতিটা irreducible factor-ই ঠিক একবার করে থাকে।

Exercise 36: $\mathbb{Q}$-এর উপরে একটা square matrix-এর characteristic polynomial হল $(1+x^2)(1-x)$. এর RCF বার করো। ■

এবার কিছু উদাহরণ দেখি, যেখানে characteristic polynomial-এর মধ্যে অন্ততঃ একটা irreducible factor একাধিকবার এসেছে।

Example 33: যদি characteristic polynomial হয় $p_1(x)p_2(x)^2$, আর minimal polynomial হয় $p_1(x)p_2(x)$, তবে rational canonical form বার করো। এখানে $p_1(x)$ আর $p_2(x)$ দুজনেই irreducible[4] বলা আছে।

SOLUTION: একটা invariant factor তো minimal polynomial-টাই, মানে $p_1(x)p_2(x)$. অন্য invariant factor-টাকে $p_2(x)$ হতেই হবে, নইলে গুণ করে characteristic polynomial হবে না। সুতরাং rational canonical form-টা হবে $\mathrm{diag}(C_{p_2(x)}, C_{p_1(x)p_2(x)})$. ■

Example 34: যদি characteristic polynomial-টা হয় $p_1(x)p_2(x)^3$, আর minimal polynomial হয় $p_1(x)p_2(x)$, তবে rational canonical form বার করো। এখানে $p_1(x)$ আর $p_2(x)$ দুটো আলাদা আলাদা irreducible polynomial.

SOLUTION: আগের মতই একটা invariant factor তো minimal polynomial-টাই, মানে $p_1(x)p_2(x)$. এছাড়া আর কটা invariant factor থাকবে এবং কী কী, সেটাই ভাববার বিষয়। সবগুলো invariant factor গুণ করে characteristic polynomial হতেই হবে। তার মানে বাকি invariant factor-গুলোর গুণ ফল $p_2(x)^2$. এমন কি হতে পারে, যে $p_2(x)^2$-টা নিজেই একটা invariant factor? না, কারণ $p_2(x)^2$ দিয়ে $p_1(x)p_2(x)$-কে ভাগ যায় না। তবে একটাই সম্ভাবনা পড়ে থাকে, দুটো invariant factor, এবং তারা দুজনেই $p_2(x)$. তার মানে rational canonical form-টা হবে $\mathrm{diag}(C_{p_2(x)}, C_{p_2(x)}, C_{p_1(x)p_2(x)})$. ■

Exercise 37: যদি characteristic polynomial-টা হয় $p_1(x)^2p_2(x)^2$, আর minimal polynomial হয় $p_1(x)p_2(x)$, তবে rational canonical form বার করো। এখানে $p_1(x)$ আর $p_2(x)$ দুটো আলাদা আলাদা irreducible

[4]কোনো field-এর উপরে একটা polynomial-কে irreducible বলা মানে ওই field-এর উপর polynomial-টাকে factorise করা যায় না।

polynomial. ■

এ পর্যন্ত খালি characteristic polynomial আর minimal polynomial জেনেই rational canonical form বার করে দেওয়া যাচ্ছিল। অবশ্যই সেটা সর্বদা হবে না[5]। এমন একটা উদাহরণ দেখাচ্ছি এবার।

Example 35: যদি characteristic polynomial হয় $p(x)^4$, আর minimal polynomial হয় $p(x)^2$, তবে rational canonical form বার করো। এখানে $p(x)$ হল irreducible.

SOLUTION: এখানে দুটো উত্তর সম্ভব--

- এক, $\mathrm{diag}(C_{p(x)^2}, C_{p(x)^2})$,
- আর দুই, $\mathrm{diag}(C_{p(x)}, C_{p(x)}, C_{p(x)^2})$.

■

Exercise 38: যদি characteristic polynomial হয় $p(x)^5$, আর minimal polynomial হয় $p(x)^2$, তবে কী কী RCF হতে পারে? এখানে $p(x)$ হল irreducible. ■

68.4 RCF থেকে rank, nullity বার করা

কোনো matrix-এর RCF দেওয়া থাকলে, তার rank এবং nullity বার করা যায়। কায়দাটা কয়েকটা অংকের মাধ্যমে দিচ্ছি।

Exercise 39: ধরো $p(x)$ কোনো একটা monic polynomial যার degree হল $n \in \mathbb{N}$. তবে দেখাও যে--

$$r(C_{p(x)}) = \begin{cases} n-1 & \text{if } p(0)=0 \\ n & \text{otherwise} \end{cases}.$$

■

Exercise 40: একটা RCF দিলাম $\mathrm{diag}(C_{p_1(x)}, ..., C_{p_n(x)})$. ধরো $p_i(x)$-দের মধ্যে ঠিক k জনের জন্য $p_i(0) = 0$ হয়। তবে দেখাও যে, RCF-টার nullity হবে k. ■

Exercise 41: আগের অংকে যদি $p_t(0) = 0$ হয়, এবং $\forall i < t \quad p_i(0) \neq 0$ হয়, তবে দেখাও যে, RCF-টার nullity হবে $k - t + 1$. ■

DAY 69
Jordan Canonical Form (part 3)

[5]কারণ তাই যদি হত, তবে দুটো matrix-এর similarity পরীক্ষা করবার জন্য এই দুটো polynomial জানাই যথেষ্ট হত, এতসব normal form-দের ডাক পড়ত না।

আমরা JCF নিয়ে এ যাবৎ যা যা দাবী করেছি, আজকে সব কিছু প্রমাণ করব। প্রমাণগুলোকে ঠিক সহজ বলা যায় না। রবীন্দ্রনাথ ঠাকুর কোনো এক প্রসঙ্গে বলেছিলেন--যে জিনিসটা পাতে পড়লে উপাদেয়, সেটাই মাথায় পড়লে গুরুতর হয়ে উঠতে পারে। একটা কাঁঠাল তোমার পাতে না দিয়ে যদি মাথা লক্ষ্য করে ছুঁড়ে মারি, তবে মাঝের থেকে মাথা ফেটে রক্ত বেরোনোই সার হবে! তা, আমাদের এই প্রমাণগুলোও কতকটা সেরকম। পুরোটা জিনিসটা একবারে তোমার মাথা লক্ষ্য করে দেগে দিলে গণিতে-শোণিতে মিশে একাকার হবে। তাই ধাপে ধাপে এগোব। যারা নিতান্ত পরীক্ষাপাশের জন্য পড়ছ, বা linear algebra-র মূল ধারণাটুকু পাওয়াই যাদের লক্ষ্য, তারা এই জায়গাটা বাদ দিয়ে যেও। বস্তুতঃ, JCF-এর প্রমাণ linear algebra শেখার জন্য ভয়ানক জরুরী কিছু নয়।
গোড়াতেই বলে নিই, কী কী দেখাতে চাই--

- প্রথমতঃ, যদি A যেকোনো একটা complex, square matrix হয়, তবে এমন একটা JCF পাওয়া যাবে J, যাতে A এবং J পরস্পরের সঙ্গে similar হয়, অর্থাৎ এমন P থাকে, যাতে $A = PJP^{-1}$ হয়। এই হল existence অংশটা।

- দ্বিতীয়তঃ uniqueness, মানে যদি দুটো JCF থাকে, যারা পরস্পরের সঙ্গে similar, তবে ওদের Jordan block-গুলো একই হবে, খালি ক্রমটা আগে পরে হয়ে যেতে পারে, এই যা।

- তৃতীয়তঃ, JCF over $\mathbb{R}$-এর প্রমাণটাও করতে হবে।

যেহেতু প্রমাণগুলো গুছিয়ে লিখতে গেলে বড্ড লম্বা, তাই আমরা প্রতি ক্ষেত্রে প্রমাণের মূল ধারণাটুকু দিয়ে ছেড়ে দেব। শুরু করব সবচেয়ে সহজ জিনিসটা দিয়ে, uniqueness-এর প্রমাণ।

69.1 Uniqueness-এর প্রমাণ

ধরো দুটো matrix আছে, A_1 আর A_2, দুজনেই JCF, এবং পরস্পরের সঙ্গে similar. তবে দেখাব যে, ওদের মধ্যে একই সব Jordan block আছে, ক্রমটা আলাদা হলেও হতে পারে। যেহেতু A_1 আর A_2 দুজনে similar, তাই ওদের eigenvalue-গুলো এবং তাদের algebraic multiplicity-ও একই। এদিকে ওরা দুজনেই triangular. তাই ওদের diagonal entry-গুলোই ওদের eigenvalue. ফলে ওদের diagonal entry-গুলো মিলে যেতে বাধ্য (ক্রমটা না মিলতে পারে)। আবার প্রতিটা eigenvalue-র geometric multiplicity-ও মিলতে বাধ্য। তাই Jordan block-এর সংখ্যাও মিলতে বাধ্য। কিন্তু Jordan block-গুলোর সাইজ কেন মিলবে, সেটা এ থেকে এখনই বোঝা যাচ্ছে না। তার জন্য আমাদের সেই বিন্দুর কায়দাটা দিয়ে ভাবব। প্রতিটা eigenvalue-র জন্য বিন্দুর কায়দাটা মনে করিয়ে দিই। ধরো A_1 নিয়ে কাজ করছি, এবং λ তার একটা eigenvalue.

> তবে প্রথমে $B_1 = A_1 - \lambda I$ নিয়ে প্রথম row-তে $\nu(B_1)$-খানা বিন্দু, প্রথম দুটো row মিলিয়ে $\nu(B_1^2)$-খানা, প্রথম তিনটে row মিলিয়ে $\nu(B_1^3)$-খানা, এইভাবে বিন্দু এঁকেছিলাম। তারপর প্রতিটা column থেকে একটা করে Jordan block পেয়েছিলাম।

তার মানে Jordan block-গুলোর সাইজ সম্পূর্ণভাবে নির্ধারিত হয়ে যাচ্ছে $\nu(B_1)$, $\nu(B_1^2)$, $\nu(B_1^3)$, ইত্যাদি থেকে। একইভাবে A_2-র বেলায় আমরা কাজ করব $B_2 = A_2 - \lambda I$ নিয়ে। এদিকে A_1 আর A_2 যেহেতু similar, তাই B_1 আর B_2-ও পরস্পরের সঙ্গে similar. সুতরাং যেকোনো k-র জন্যই $\nu(B_1^k) = \nu(B_2^k)$ হবে। ফলে A_1 আর A_2-র Jordan block-দের সাইজও সমান হতে বাধ্য। ব্যস্! খুব কঠিন নয় যুক্তিটা, কী বলো?

69.2 JCF over $\mathbb{R}$

আমরা real field-এর উপরে JCF-এর একটা সংস্করণের কথা বলেছিলাম। সেটা এমনি JCF থেকে (মানে $\mathbb{C}$-এর উপর JCF থেকে) কী করে আসে বলি। ধরো $A_{n\times n}$ একটা real matrix, যার একটা eigenvalue হল $\alpha = \lambda + i\mu$, যেখানে $\lambda, \mu \in \mathbb{R}$ এবং $\mu \neq 0$. এর জন্য কোনো একটা eigenvector নেওয়া যাক, $\vec{z} \in \mathbb{C}^n$. আমরা $\vec{z} = \vec{u} + i\vec{v}$ লিখব, যেখানে $\vec{u}, \vec{v} \in \mathbb{R}^n$. আমাদের কাছে ছিল

$$A\vec{z} = w\vec{z}. \qquad (*)$$

আমরা একে বড় করে লিখতে পারি

$$A(\vec{u} + i\vec{v}) = (\lambda + i\mu)(\vec{u} + i\vec{v}).$$

এর real এবং imaginary অংশ আলাদা করে লিখলে হয়--

$$\begin{aligned} A\vec{u} &= \lambda\vec{u} - \mu\vec{v}, \\ A\vec{v} &= \mu\vec{u} + \lambda\vec{v}. \end{aligned}$$

এটাকে matrix-এর আকারে লেখা যায় এইভাবে--

$$A\begin{bmatrix} \vec{u} & \vec{v} \end{bmatrix} = \begin{bmatrix} \vec{u} & \vec{v} \end{bmatrix} \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}. \qquad (**)$$

একবার (*)-এর সঙ্গে এই (**)-কে তুলনা করলেই দেখবে কীভাবে non-real w-র জায়গায় $\begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}$ এসে গেল। ঠিক যেন w-টা ছিল একটা 1×1 Jordan block, $J_1(w)$. সেটার জায়গায় এল $R_1(w)$. একইরকম ব্যাপার যেকোনো সাইজের Jordan block-এর জন্যই হবে। নীচের অংকে দেখাতে বলেছে যে, $J_2(w)$-এর জায়গায় আসবে $R_2(w)$.

Exercise 42: ধরো $A_{n\times n}$ একটা real matrix যার জন্য

$$A\begin{bmatrix} \vec{z}_1 & \vec{z}_2 \end{bmatrix} = \begin{bmatrix} \vec{z}_1 & \vec{z}_2 \end{bmatrix} J_2(w)$$

হয়। এর real আর imaginary অংশ আলাদা করে দেখাও যে, যদি $\vec{z}_1 = \vec{u}_1 + i\vec{v}_1$ এবং $\vec{z}_2 = \vec{u}_2 + i\vec{v}_2$ হয়, যেখানে $\vec{u}_1, \vec{v}_1, \vec{u}_2, \vec{v}_2 \in \mathbb{R}^n$, তবে

$$A\begin{bmatrix} \vec{u}_1 & \vec{v}_1 & \vec{u}_2 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} \vec{u}_1 & \vec{v}_1 & \vec{u}_2 & \vec{v}_2 \end{bmatrix} R_2(w)$$

হবে। ■

General প্রমাণটা গুছিয়ে লিখতে গেলে notation-গুলো বিচ্ছিরি হয়ে যাবে। তাই আমরা এখানে তার মধ্যে যাচ্ছি না।

69.3 Existence-এর প্রমাণ

এইবার আমরা সবচেয়ে শক্ত জিনিসটায় হাত দেব, JCF-এর existence-এর প্রমাণ। কোনো একটা complex, square matrix নিয়ে শুরু করব, ধরো A. দেখাব যে, $A = PJP^{-1}$ লেখা যায় যেখানে P হল nonsingular এবং J হল JCF.

আমরা JCF-এর J-টা বার করার জন্য একটা কায়দা বলেছিলাম বিন্দু এঁকে এঁকে। যদি প্রতিটা বিন্দুর জায়গায় একটা করে উপযুক্ত vector লিখি, তবে তাদেরকে পাশাপাশি column হিসেবে দাঁড় করিয়ে দিলেই P-ও তৈরী হয়ে যাবে। প্রথম ধাপে বলব কীভাবে সেই উপযুক্ত vector-দের নির্বাচন করতে হবে। তার পরের ধাপে দেখাব যে, সেই vector-গুলো linearly independent হবে। তৃতীয় ধাপে দেখাব যে ওরা $\mathbb{C}^n$-এর একটা basis হবে (ফলে ওদের পাশাপাশি রেখে matrix বানালে সেটা nonsingular হবে)।

69.3.1 প্রথম ধাপ-- vector নির্বাচন করা

বিন্দুর কায়দাটায় আমরা প্রতিটা eigenvalue নিয়ে আলাদা করে একই কাজ করেছিলাম। এখানেও vector নির্বাচনের সময়ে আমরা প্রতিটা eigenvalue নিয়ে আলাদা করে কাজ করব। বোঝানোর সুবিধার জন্য ধরো আমাদের matrix-টার খালি একটাই eigenvalue আছে, λ. যদি matrix-টার সাইজ হয় $n \times n$, তবে λ-র algebraic multiplicity-ও হবে n. আগের মত এখানেও আমরা $B = A - \lambda I$ নিয়ে কাজ করব। লক্ষ্য করো যে--

1. $\mathscr{C}(B^0) = \mathbb{C}^n$ হবেই। $\mathscr{C}(B)$ হবে তার মধ্যে একটা subspace. আবার $\mathscr{C}(B^2)$ হচ্ছে তারও ভিতরে আরেকটা subspace.
2. একইভাবে যেকোনো B-এর জন্যই $\mathscr{N}(B^0) = \{\vec{0}\}$, তার চেয়ে বড় হবে $\mathscr{N}(B)$, তারও চেয়ে বড় হবে $\mathscr{N}(B^2)$, ইত্যাদি।

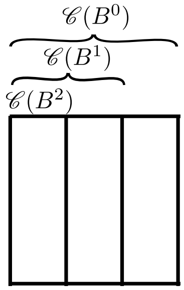


Fig 11

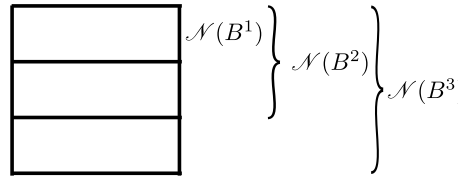


Fig 12

এ দুটো জিনিস যেকোনো B-এর জন্যই হবে। কিন্তু আমাদের বিশেষ B-এর জন্য আরেকটু বেশী কিছু বলা যায়। লক্ষ্য করো যে, B-এর সব eigenvalue-ই 0, এবং তাই[6] B হল nilpotent.
তার মানে কোনো একটা $k \in \mathbb{N}$-এর জন্য $B^k = O$ হবে। উদাহরণ হিসেবে ধরো $B^3 = O$. তবে $\mathscr{C}(B^3) = \{\vec{0}\}$ এবং $\mathscr{N}(B^3) = \mathbb{C}^3$. তার মানে সব মিলিয়ে হল--

$$\{\vec{0}\} = \mathscr{C}(B^3) \subseteq \mathscr{C}(B^2) \subseteq \mathscr{C}(B^1) \subseteq \mathscr{C}(B^0) = \mathbb{C}^n$$

এবং

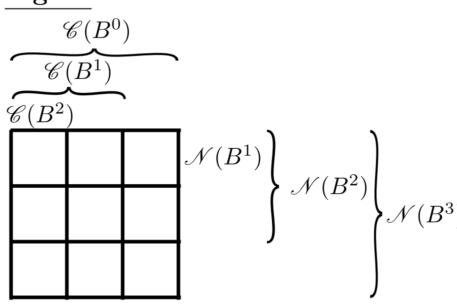
$$\{\vec{0}\} = \mathscr{N}(B^0) \subseteq \mathscr{N}(B^1) \subseteq \mathscr{N}(B^2) \subseteq \mathscr{N}(B^3) = \mathbb{C}^n.$$

পুরো ব্যাপারটাকে একটা ছকের মত করে লিখব। ছকটা আমাদের খুব কাজে লাগবে, তাই ভালো করে বুঝে বুঝে পড়। Fig 11-এ পুরো $\mathbb{C}^n$-কে একটা square-এর মত করে এঁকেছি। এটাকে তিনটে column-এ ভেঙেছি। প্রথম column-টা হল $\mathscr{C}(B^2)$, প্রথম দুটো column মিলে $\mathscr{C}(B)$. আর যদি তিনটে column-ই নাও, তবে পাবে $\mathscr{C}(B^0)$, মানে $\mathbb{C}^n$. লক্ষ্য করো, দ্বিতীয় column-টা হল সেই সব vector-এর set, যারা $\mathscr{C}(B)$-তে আছে, কিন্তু $\mathscr{C}(B^2)$-এ নেই। এই গেল column space-এর বিচারে ভাগ। এবার Fig 12-এ সেই একই square-কে তিনটে row-তে ভেঙেছি nullspace-এর বিচারে। প্রথম row-টা $\mathscr{N}(B)$, প্রথম দুটো row মিলে $\mathscr{N}(B^2)$, এইরকম। সুতরাং খালি দ্বিতীয় row-টা হল সেই সব vector-এর set, যারা $\mathscr{N}(B^2)$-এ আছে, কিন্তু $\mathscr{N}(B)$-এ নেই। যদি সবগুলো row একসঙ্গে নাও, তবে পাবে $\mathscr{N}(B^3)$, যেটা সম্পূর্ণ $\mathbb{C}^n$. এবার Fig 11 আর Fig 12-কে একই ছবিতে বসালে পাবে Fig 13, যেটা আমাদের প্রচুর কাজে লাগবে। আপাততঃ নীচের অংক কয়েকটা করে ছকটার সঙ্গে বন্ধুত্ব করে নাও।

Example 36: আমরা $B^3 = O$ নিয়েছি। তাই $B^2\vec{x}$-জাতীয় কোনো vector-কে B দিয়ে গুণ করলে $\vec{0}$ পাবে, কারণ $B \times B^2\vec{x} = B^3\vec{x} = \vec{0}$. তার মানে $\mathscr{C}(B^2) \subseteq \mathscr{N}(B)$. এ থেকেই বলা যায় যে, ছকটার দুটো ঘর আসলে ফাঁকা। কোন্ দুটো ঘর?

[6] এই ব্যাপারটা আমরা সপ্তম অধ্যায়ে প্রমাণ করেছিলাম, Schür decomposition-এর উপকারিতা আলোচনা করার সময়ে।

Fig 13



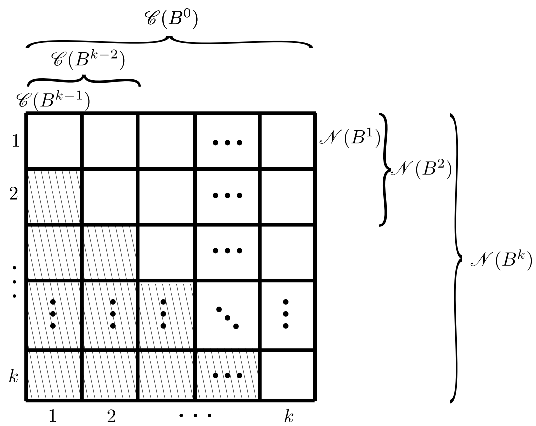


Fig 14

SOLUTION: প্রথম column-এর নীচের দুটো ঘর, মানে matrix হিসেবে ভাবলে $(2,1)$ আর $(3,1)$ ঘর দুটো। ওরা $\mathscr{C}(B^2)$-এর মধ্যে আছে, অথচ $\mathscr{N}(B)$-এর মধ্যে নেই। এর ফলে ছকের $(2,1)$ আর $(3,1)$ ঘর দুটো ফাঁকা হয়ে গেল। ■

Exercise 43: আগের অংকের মত এটাও দেখাও যে, $\mathscr{C}(B) \subseteq \mathscr{N}(B^2)$. এ থেকে কোন্ কোন্ ঘরকে ফাঁকা বলা যায়? ■

এখানে আমরা $B^3 = O$ নিয়ে কাজ করছিলাম। যদি $B^k = O$ হত, তবে ছবিটা হত Fig 14-এর মত। এখানে আমরা পুরো $\mathbb{C}^n$-কে অনেকগুলো subset-এ partition করেছি। প্রত্যেকটা ঘর হল একটা subset. এদের মধ্যে শেড করা ঘরগুলো ফাঁকা।

Exercise 44: যদি $k = 4$ নিই, মানে $B^4 = O$ হয়, তবে ছকটা হবে 4×4 সাইজের। সেটা আঁকো, লেবেল করো, এবং শূন্য ঘরগুলোকে শেড করে দাও।
■

Exercise 45: Fig 14-এর কোন্ ঘরে $\vec{0}$ আছে? ■

আমরা কাজ করব খালি nonnull vector-দের নিয়ে। লক্ষ্য করো--

- কোনো $\vec{x} \neq \vec{0}$ যদি i-th row-তে থাকে, তার মানে $\vec{x} \in \mathscr{N}(B^i)$ কিন্তু $\vec{x} \notin \mathscr{N}(B^{i-1})$.
- কোনো $\vec{x} \neq \vec{0}$ যদি j-th column-এ থাকে, তার মানে $\vec{x} \in \mathscr{C}(B^{k-j})$ কিন্তু $\vec{x} \notin \mathscr{C}(B^{k-j+1})$.

আগে আমরা $\nu(B)$-সংখ্যক বিন্দু পাশাপাশি লিখে শুরু করেছিলাম। এবার প্রতিটা বিন্দুর জায়গায় একটা করে vector নেব, যাতে সবগুলো মিলে $\mathscr{N}(B)$-র একটা basis হয়। যে সে basis হলে হবে না, একটা বিশেষ কায়দায় নিতে হবে। প্রথমে $\mathscr{N}(B) \cap \mathscr{C}(B^{k-1})$-এর একটা basis নিয়ে শুরু করতে হবে। এই vector-গুলো সবাই থাকবে $(1,1)$ ঘরে। সেটাকে extend করতে হবে $(1,2)$ ঘরের কিছু vector দিয়ে, যাতে পুরোটা মিলে $\mathscr{N}(B) \cap \mathscr{C}(B^{k-2})$-এর একটা basis হয়। এইভাবে চলতে চলতে সব শেষে জুড়ব $(1,k)$ ঘরের কিছু vector, যাতে সব মিলে $\mathscr{N}(B)$-এর basis হয়। মনে রেখো $\mathscr{N}(B)$ হল আমাদের ছকের পুরো প্রথম row-টা। এবার আমরা এই vector-গুলোকে "কোণাকুণি ডানদিকে একঘর করে নামাব"। এই ব্যাপারটা একটু বুঝে নেওয়া যাক। Fig 15-এ কোণাকুণি গতিটাকে তীরচিহ্ন দিয়ে দেখিয়েছি। ধরো, (i,j) ঘরে

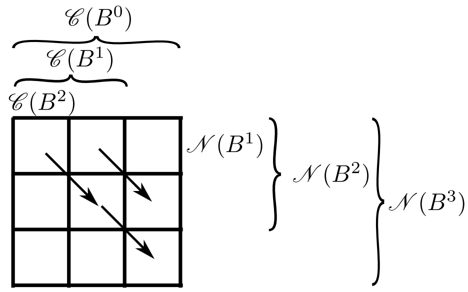


Fig 15

কোনো $\vec{x} \neq \vec{0}$ নিয়ে শুরু করলাম, যেখানে $i, j < k$. তবে তার জন্য $(i+1, j+1)$ ঘরে এমন একটা[7] $\vec{y}$ থাকবেই, যাতে $\vec{x} = B\vec{y}$ হয়। যদি জানতে চাও, কেন থাকবে, তবে নীচের অংকটা করো। খুব একটা শক্ত কিছু নয়, তবে প্রথমবার পড়ার সময়ে এই অংকটা বাদ দিয়ে যেতে পারো। আমরা এইরকম $\vec{y}$-কে বলব $\vec{x}$-এর একটা "সন্তান", আর $\vec{x}$-কে বলব $\vec{y}$-এর "বাবা"[8] মনে রেখো যে, কোনো ঘরের vector-দের সন্তানরা থাকে তার নীচের ডানদিকের ঘরে। যদি $i = k$ বা $j = k$ হয়ে যায়, তবে তার নীচে ডানদিকে কোনো ঘর নেই, তাই তার সন্তানও থাকে না। আরেকটা জিনিসও লক্ষ্য করো--কোনো সন্তানকে B দিয়ে গুণ করলে তার বাবা হয়ে যায় (মানে, B for Baba, আর কি!)!

Exercise 46: ধরো, $\vec{x} \neq \vec{0}$ রয়েছে (i, j) ঘরে, যেখানে $i, j < k$. তবে দেখাও যে, $(i+1, j+1)$ ঘরে এমন একটা $\vec{y}$ থাকবেই, যাতে $\vec{x} = B\vec{y}$ হয়।
HINT:
এখানে $\vec{x} \in \mathscr{C}(B^{k-j})$, কিন্তু $\vec{x} \notin \mathscr{C}(B^{k-j+1})$. আবার $\vec{x} \in \mathscr{N}(B^i)$, কিন্তু $\vec{x} \notin \mathscr{N}(B^{i-1})$. যেহেতু $\vec{x} \in \mathscr{C}(B^{k-j})$, তাই কোনো একটা $\vec{z}$-এর জন্য $\vec{x} = B^{k-j}\vec{z}$ হবে। তুমি $\vec{y} = B^{k-j-1}\vec{z}$ নাও। দেখাও যে, $\vec{y}$ থাকবে $(i+1, j+1)$ ঘরে।
■

আমরা এক্ষুণি $\mathscr{N}(B)$-এর যে basis-টা বানালাম, এবার সেই vector-গুলোর সন্তানদের নেব। যেমন ধরো $(1, 1)$ ঘরের vector-গুলোর সন্তানরা থাকবে $(2, 2)$ ঘরে, তাদের সন্তানরা থাকবে $(3, 3)$ ঘরে, এইরকমভাবে নাতিপুতি হতে হতে এসে পৌঁছবে (k, k) ঘর পর্যন্ত। একইভাবে $(1, 2)$ ঘরের vector-দের সন্তান-সন্ততি, নাতিপুতি বিস্তৃত হবে $(k-1, k)$ ঘর পর্যন্ত। এইভাবে তুমি অনেকগুলো vector পাবে। এখানেই প্রথম ধাপের শেষ।
দ্বিতীয় ধাপে দেখাব যে এই সব vector মিলে একটা linearly independent set হয়। এবং তার পরের ধাপে দেখাব যে, এরা পুরো $\mathbb{C}^n$-কে span করে। তার মানে সব মিলিয়ে এই vector-গুলো $\mathbb{C}^n$-এর একটা basis দিচ্ছে। এই basis-এর vector-গুলোকে পাশাপাশি column হিসেবে বসিয়েই তৈরী হবে আমাদের P.

69.3.2 দ্বিতীয় ধাপ-- linearly independent দেখানো

এই vector-গুলো মিলে যে একটা linearly independent set হয়, তার প্রমাণটা গুছিয়ে লেখার পক্ষে একটু জটিল হলেও ছবি দিয়ে বোঝার পক্ষে বেশ মজার। আমরা আবার Fig 15-এ দেখানো ছকটাতে ফিরে আসি। ঘরগুলোতে কয়েকটা করে vector রয়েছে। আমাদের দেখাতে হবে সবগুলো মিলে একটা linearly independent set হয়। এর মধ্যে একটা জিনিস আমাদের হাতে আছে-- প্রথম row-এর সব vector-ই linearly independent, কারণ আমরা সেভাবেই ওদের তৈরী করেছি। ওটাই বার বার করে ব্যবহার করব।
তার জন্য ওদের সবার একটা linear combination নিয়ে শূন্যর সমান করে দেওয়া যাক। আমরা দেখাব যে, সেক্ষেত্রে প্রতিটা coefficient-ই শূন্য হতে বাধ্য। এই linear combination-টাকে কয়েকটা গুচ্ছে ভেঙে কল্পনা করো, যেখানে একেকটা গুচ্ছ হল একেকটা ঘরের vector-গুলো নিয়ে। উদাহরণস্বরূপ $k = 3$-এর জন্য ব্যাপারটা এইরকম--

$$\text{প্রথম row-এর তিনটে গুচ্ছ} + \text{দ্বিতীয় row-এর দুটো গুচ্ছ} + \text{তৃতীয় row-এর একটা গুচ্ছ} = \vec{0}.$$

[7]"একটা" মানে "অন্ততঃ একটা", কোনো uniqueness দাবী করছি না।
[8]চাইলে "মা"-ও বলতে পারো।

এইবার দুপাশকেই B দিয়ে গুণ করে দাও, বাঁদিক থেকে। এর ফলে প্রথম row-এর গুচ্ছ তিনটে পটল তুলবে, কারণ ওরা সবাই $\mathscr{N}(B)$-এর সদস্য। দ্বিতীয় row-এর গুচ্ছ দুটো ওদের বাবায় পরিণত হবে, ফলে প্রথম row-তে উঠবে। একইভাবে তৃতীয় row-এর গুচ্ছটার পদোন্নতি হবে দ্বিতীয় row-তে। B দিয়ে গুণ করার ফলে সন্তানের বাবা হয়ে ওঠাটাই আমাদের যুক্তির মোক্ষম ধাপ।
এইবার ফের আরেকবার B দিয়ে বাঁদিক থেকে গুণ করো (মানে সব মিলিয়ে B^2 দিয়ে গুণ করা হল)। এর ফলে যে দুটো গুচ্ছ আগের ধাপে প্রথম row-এ উঠেছিল, তাদের জীবনাবসান হবে, আর যেটা আগের ধাপে তৃতীয় row থেকে দ্বিতীয় row-তে উঠেছিল, সে এখন ফের কোণাকুণি পথে উঠে প্রথম row-তে হাজির হবে। ব্যস্, এখন খালি একটা গুচ্ছ পড়ে আছে, এবং সেটা আছে প্রথম row-এ। প্রথম row-এর গুচ্ছরা যে linearly independent সে তো আমরা জানিই। সুতরাং ওই গুচ্ছের যাবতীয় coefficient শূন্য হয়ে যেতে বাধ্য। তার মানে একটা গুচ্ছের coefficient-দের বিদায় করা গেল। এবার ফের গোড়ার জিনিসটায় ফিরে যাই, খালি শেষের গুচ্ছটা আর লেখার দরকার রইল না--

প্রথম row-এর তিনটে গুচ্ছ + দ্বিতীয় row-এর দুটো গুচ্ছ $= \vec{0}$.

এবার ফের B দিয়ে গুণ করব বাঁদিক থেকে। প্রথম row-এর গুচ্ছ তিনটে লোপ পাবে, আর দ্বিতীয় row-এর গুচ্ছগুলো প্রথম row-এ এসে হাজির হবে। প্রথম row-এর vector-রা তো linearly independent আছেই (সেভাবেই ওদের নেওয়া হয়েছিল)। অতএব ওদের সবার coefficient-ও শূন্য হতে বাধ্য।
এবার যদি গোড়ার জিনিসটার দিকে আবার তাকাই, তবে সেখানে পড়ে আছে খালি--

প্রথম row-এর তিনটে গুচ্ছ $= \vec{0}$.

যেহেতু প্রথম row-এর সব vector মিলে একটা linearly independent হয়, তাই ওদের সবার coefficient-ই শূন্য হতে বাধ্য। ব্যস্, প্রমাণ শেষ!

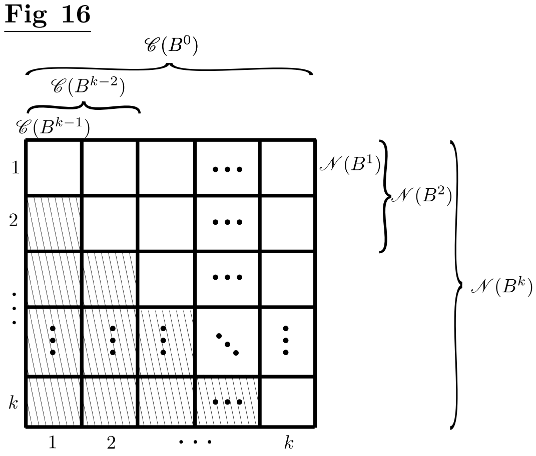
69.3.3 তৃতীয় ধাপ-- spanning দেখানো

এই ধাপে আমাদের কাজ হল এইটা দেখানো যে, সবগুলো vector মিলে $\mathbb{C}^n$-কে span করে। আমরা তার চেয়ে আরো জোরালো কিছু দেখাব। আমরা দেখাব যে--

যেকোনো $p \in \{1, ..., k\}$-র জন্যই প্রথম p-খানা row-এর সব vector মিলে $\mathscr{N}(B^p)$-এর একটা basis হয়।

তবে $p = k$ বসালেই পাবো যে, সবগুলো vector মিলে $\mathscr{N}(B^k)$-কে মানে পুরো $\mathbb{C}^n$-কে span করে।
এখানেও যুক্তিটা ছক দিয়ে বুঝতে সহজ। মূল যুক্তিটা হল p-র উপর induction, মানে $p = 1, 2, 3, ...$ করে এগোনো। যদি $p = 1$ হয়, তবে দেখানোর কিছু নেই, কারণ প্রথম row-এর vector গুলোকে তো আমরা $\mathscr{N}(B)$-এর একটা basis হিসেবেই নিয়েছি। এর পরের ধাপে $p = 2$. সেটার জন্য প্রমাণটা করে দেখাচ্ছি। সেটার যুক্তিটা বুঝলেই, বাকি সব p-এর

Fig 16



জন্যও একইভাবে হবে। প্রথমে Fig 16 দেখে নাও। এই ছবিটা আগেও একবার দেখেছ। চট্ করে দেখার সুবিধার জন্য এখানে আরেকবার দিয়েছি। আমরা দেখাতে চাইছি যে, প্রথম আর দ্বিতীয় row-এর vector-গুলো $\mathscr{N}(B^2)$-কে span করে, অর্থাৎ $\mathscr{N}(B^2)$-এর যেকোনো vector-কেই ওদের linear combination হিসেবে লেখা যায়। যেকোনো একটা $\vec{x} \in \mathscr{N}(B^2)$ নেওয়া যাক। তাহলে $B^2\vec{x} = \vec{0}$. সুতরাং $B\vec{x} \in \mathscr{N}(B)$. আবার $B\vec{x} \in \mathscr{C}(B)$-ও বটে। সুতরাং $B\vec{x} \in \mathscr{C}(B) \cap \mathscr{N}(B)$. অতএব $B\vec{x}$-কে $(1,1), ..., (1, k-1)$ ঘরের vector-গুলোর linear combination হিসেবে লেখা যাবে। এবার মনে করে দেখো, এরা হল দ্বিতীয় row-এর vector-গুলোর বাবা। সুতরাং

$$B\vec{x} = B(\text{দ্বিতীয় row-এর কোনো linear combination}).$$

তার মানে--

$$B(\vec{x} - (\text{দ্বিতীয় row-এর কোনো linear combination})) = \vec{0}.$$

তার মানে ওই $(\vec{x} - (\text{দ্বিতীয় row-এর কোনো linear combination}))$-টা রয়েছে $\mathscr{N}(B)$-এর মধ্যে। সুতরাং ওটাকে প্রথম row-এর linear combination হিসেবে লেখা যাবে। সুতরাং সব মিলিয়ে পাচ্ছি যে, $\vec{x}$-কে প্রথম আর দ্বিতীয় row-এর linear combination হিসেবে লেখা যাচ্ছে, ঠিক যেমনটা আমাদের দরকার ছিল।

Exercise 47: একই যুক্তিতে $p = 3$-র জন্য প্রমাণটা করো। ■

আজকের পড়া তবে এই পর্যন্তই। প্রমাণের প্রধান অংশটা পার হয়েছে। এখনও কিছু প্যাঁচ অবশ্য বাকি আছে। সেগুলো কালকের জন্য তোলা থাক।

DAY 70
Jordan Canonical Form (part 4)

70.1 Existence-এর প্রমাণ চলছে

গতকাল প্রথম তিনটে ধাপ করেছি। আজকে পরবর্তী আকর্ষণ।

70.1.1 চতুর্থ ধাপ--একটা eigenvalue-র জন্য Jordan block তৈরী

এই যে আমরা $\mathbb{C}^n$-এর একটা basis পেলাম (মানে সবগুলো ঘরের vector মিলিয়ে), তাদের পরপর column-এর মত বসিয়ে আমাদের P-টা বানাব। কিন্তু পরপর বসানোর সময়ে একটা বিশেষ ক্রমে কাজটা করতে হবে। সেটা হল প্রথমে $(1,1)$ ঘরের কোনো একটা vector থেকে শুরু করে, তার সন্তান, নাতিপুতি ইত্যাদি পরপর লেখা শেষ করে তারপর $(1,1)$ ঘরের আরেকটা কোনো vector-এ হাত দেওয়া। এইভাবে $(1,1)$ ঘরের সব vector শেষ হলে $(1,2)$-কে নিয়ে একই কাজ করা, ইত্যাদি।

Example 37: Fig 17-এ যে vector-গুলো দেখিয়েছি, তাদেরকে কার পর কাকে সাজাবে। এখানে তীরচিহ্নগুলো বলছে কে কার সন্তান।

SOLUTION: একভাবে হতে পারে এরকম--

$$\underbrace{\vec{a},\quad \vec{f},\quad \vec{i},}\quad \underbrace{\vec{b},\quad \vec{g},\quad \vec{j},}\quad \underbrace{\vec{c},\quad \vec{h},}\quad \vec{d},\quad \vec{e}.$$

তলায় ঢেউ-খেলানো ব্র্যাকেট দিয়ে একেকটা বংশ দেখিয়েছি, শেষের $\vec{d}$ আর $\vec{e}$ নিজেরাই একেকটা বংশ। এদের মধ্যে চাইলে প্রথম দুটো বংশ আগে পরে করা চলত, কারণ ওদের দুজনেরই আদি পূর্বপুরুষ একই ঘরের। একই কারণে $\vec{d}$ আর $\vec{e}$-এর ক্রমও ওল্টানো চলত। ■

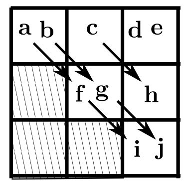


Fig 17

এইভাবে vector-গুলোকে পরপর column হিসেবে লিখে যে matrix-টা পাবে, সেটাকেই P নেব। প্রতিটা বংশ থেকে একেকটা Jordan block তৈরী হবে। সেটা কী করে হচ্ছে, বুঝে নেওয়া যাক।

Example 38: Fig 17-এ একটা বংশ হল $\vec{a}$, $\vec{f}$, $\vec{i}$. এ থেকে কীভাবে একটা Jordan block আসছে বলো।

SOLUTION: এখানে

$$\begin{aligned} B\vec{a} &= \vec{0}, \\ B\vec{f} &= \vec{a}, \\ B\vec{i} &= \vec{f}. \end{aligned}$$

যেহেতু $B = A - \lambda I$, তাই--

$$\begin{aligned} A\vec{a} &= \lambda\vec{a}, \\ A\vec{f} &= \vec{a} + \lambda\vec{f}, \\ A\vec{i} &= \vec{f} + \lambda\vec{i}. \end{aligned}$$

এটাকেই matrix-এর আকারে লিখলে হয়--

$$A\begin{bmatrix} \vec{a} & \vec{f} & \vec{i} \end{bmatrix} = \begin{bmatrix} \vec{a} & \vec{f} & \vec{i} \end{bmatrix}\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

দেখলে, Jordan block-টা কীভাবে জন্মাল? ■

Exercise 48: একইভাবে $\vec{c}$, $\vec{h}$ বংশটা থেকে একটা Jordan block বানাও, J, যাতে

$$A\begin{bmatrix} \vec{c} & \vec{h} \end{bmatrix} = \begin{bmatrix} \vec{c} & \vec{h} \end{bmatrix} J$$

হয়। ■

Exercise 49: উপরের দুটো অংকে আমরা দুটো আলাদা বংশ নিয়ে কাজ করেছি। দুটো মিলিয়ে এমন একটা JCF বানাও, D, যাতে

$$A\begin{bmatrix} \vec{a} & \vec{f} & \vec{i} & \vec{c} & \vec{h} \end{bmatrix} = \begin{bmatrix} \vec{a} & \vec{f} & \vec{i} & \vec{c} & \vec{h} \end{bmatrix} D$$

হয়। ■

70.1.2 পঞ্চম ধাপ--একাধিক eigenvalue

এতক্ষণ আমরা এমন একটা matrix নিয়ে কাজ করছিলাম, যার খালি একটাই eigenvalue. সেটাই n বার ফিরে ফিরে এসেছে, মানে যার algebraic multiplicity ছিল n. কিন্তু অনেক সময়েই আমাদের এমন সব $n\times n$ সাইজের matrix নিয়ে কাজ করতে হবে, যাদের একাধিক eigenvalue আছে, ধরো $\lambda_1\neq\cdots\neq\lambda_k\in\mathbb{C}$ যেখানে λ_i-এর algebraic multiplicity হল n_i. অবশ্যই $\sum n_i=n$ হবেই। এক্ষেত্রেও একই কায়দা লাগানো সম্ভব, প্রতিটা eigenvalue-র জন্য আলাদা করে। শেষে সবকিছু জুড়ে দিলেই চলবে। এই আলাদা করে করা এবং জুড়ে দেওয়ার পিছনে গুরুত্বপূর্ণ ভূমিকা পালন করবে **primary decomposition theorem** নামের একটা গুরুত্বপূর্ণ theorem. তার জন্য প্রথমে একটা নতুন সংজ্ঞা লাগবে--

> **DEFINITION: Generalised eigenspace**
>
> Let $A_{n\times n}$ be any matrix with an eigenvalue λ. Then $\mathscr{N}\big((A-\lambda I)^n\big)$ is called the **generalised eigenspace** of A corresponding to λ. The nonzero vectors in it are called **generalised eigenvectors** of A corresponding to λ.

এই সংজ্ঞাটা কেন এল, সেটা বোঝা কঠিন নয়। যখন একটাই eigenvalue থাকছিল, তখন আমরা $\mathscr{N}(A-\lambda I)$, $\mathscr{N}\big((A-\lambda I)^2\big),\ldots$ ইত্যাদি করতে করতে $\mathscr{N}\big((A-\lambda I)^n\big)$ অব্ধি গিয়েছিলাম। সেখানে $(A-\lambda I)^n=O$ হচ্ছিল, তাই $\mathscr{N}\big((A-\lambda I)^n\big)=\mathbb{C}^n$ ছিল। এখানে সেটা আর $\mathbb{C}^n$ হতে পারবে না (কারণ বাকি eigenvalue-গুলোও আছে)। সেই কারণে $\mathscr{N}\big((A-\lambda I)^n\big)$-এর একটা নতুন নাম দেওয়া দরকার পড়ল।

এই বার primary decomposition theorem-টা বলার সময় এসেছে। এর মূল কথাটা হল এই যে, যদি বিভিন্ন generalised eigenspace নাও, এবং প্রত্যেকটার একটা করে basis বার করো, তবে সেই basis-গুলোকে একত্র করলেই পুরো $\mathbb{C}^n$-এর একটা basis পেয়ে যাবে! এমন চমৎকার একটা ব্যাপার হলে আমাদের কী লাভ, সেটা নিশ্চয়ই বুঝতে পারছ। আগের ধাপ পর্যন্ত যে কায়দাটা শিখলাম, সেটা প্রতিটা eigenvalue-র বেলায় আলাদা করে লাগালে প্রতিটা generalised eigenspace-এর একটা করে basis পাবে, যেখান থেকে সেই eigenvalue-এর Jordan block-গুলো আসবে। এবার সবগুলো basis-কে পাশাপাশি বসালেই তো $\mathbb{C}^n$-এর একটা basis পাবে, সেটা থেকেই আসবে আমাদের P. আর ওদিকে বিভিন্ন eigenvalue-র Jordan block-গুলো মিলে তৈরী হবে JCF-টা।

JCF-এর existence-এর রহস্যটা এটাই। অবশ্য primary decomposition theorem-টা আমরা এখনও প্রমাণ করি নি। যাদের এখনো দম আছে, তারা সেই প্রমাণটা পরের ধাপে পাবে। যদিও প্রথমবার পড়ার সময়ে বোধহয় অতটা চাপ একবারে না নেওয়াই ভালো!

70.1.3 ষষ্ঠ ধাপ-- Primary decomposition theorem

Primary decomposition theorem-টা প্রমাণ করার জন্য প্রথমে কিছু ছোটো ছোটো theorem দিয়ে শুরু করি, যেগুলো generalised eigenspace-গুলোর বিভিন্ন প্রয়োজনীয় ধর্মের সঙ্গে আমাদের পরিচয় করাবে। প্রথমে theorem-গুলো খালি পড়ে যাও, প্রমাণ নিয়ে মাথা ঘামিও না। এগুলো থেকে primary decomposition theorem কীভাবে আসে, সেটা বুঝে যাবার পর, ধীরেসুস্থে এদের প্রমাণগুলো শিখে নিতে পারবে।

> **THEOREM**
>
> Let $A_{n\times n}$ be a matrix with an eigenvalue λ with algebraic multiplicity p and corresponding generalised eigenspace E. Then $\forall\vec{x}\in E\quad A\vec{x}\in E$.

সপ্তম অধ্যায়ে আমরা invariant subspace-এর কথা শিখেছিলাম মনে আছে? এমন একটা subspace, যার মধ্যে কোনো $\vec{x}$ নিলে পরে $A\vec{x}$-ও সেই subspace-এর মধ্যেই থাকতে বাধ্য। সেই ভাষায় বললে এই theorem-টা বলছে যে, generalised eigenspace-রা invariant subspace হয়।

Proof: দেওয়া আছে $(A-\lambda I)^n\vec{x}=\vec{0}$. দেখাতে হবে $(A-\lambda I)^n A\vec{x}=\vec{0}$. এর জন্য লক্ষ্য করো যে, $(A-\lambda I)^n$ এবং A পরস্পরের সঙ্গে commute করে, অর্থাৎ $(A-\lambda I)^n A = A(A-\lambda I)^n$ হয়। তাই $(A-\lambda I)^n A\vec{x} = A(A-\lambda I)^n\vec{x}=\vec{0}$. **[Q.E.D]**

এই কায়দাটা বেশ শক্তিশালী। A-র যেকোনো polynomial-ই A-র অন্য যেকোনো polynomial-এর সঙ্গে commute করে। ফলে যদি কোনো generalised eigenspace-এর কোনো vector-কে তুমি A-র যেকোনো polynomial দিয়ে গুণ করো, তবে গুণফলটাও ফের ওই generalised eigenspace-এর মধ্যেই থাকতে বাধ্য। এই কথাটা শীঘ্রই আমাদের কাজে লাগবে।

THEOREM

Let $A_{n\times n}$ be a matrix with an eigenvalue λ with algebraic multiplicity p and corresponding generalised eigenspace, E. Then $\dim(E)=p$.

Proof: এখানে আমরা Schür decomposition ব্যবহার করে প্রথমে matrix-টাকে একটা upper triangular matrix ধরে নেব, পরিণত করে নেব, যার প্রথম p-খানা diagonal entry হবে λ. তবে $A-\lambda I$ হবে এরকম দেখতে--

$$A-\lambda I = \begin{bmatrix} Q & * \\ O & S \end{bmatrix},$$

যেখানে $Q_{p\times p}$ এমন একটা upper triangular matrix, যার diagonal entry-গুলো সবই 0. ওদিকে $S_{n-p\times n-p}$ এমন একটা upper triangular matrix, যার diagonal entry-গুলো সকলেই non-zero. সুতরাং

$$(A-\lambda I)^n = \begin{bmatrix} Q^n & * \\ O & S^n \end{bmatrix} = \begin{bmatrix} O & * \\ O & S^n \end{bmatrix}.$$

যেহেতু এর প্রথম p-খানা column-ই হল $\vec{0}$, তাই এর nullity অবশ্যই $\geq p$. এদিকে যেহেতু $\text{gm}(\lambda,A)\leq \text{am}(\lambda,A)$, তাই nullity-টা আবার $\leq p$-ও বটে। সুতরাং দুইয়ে মিলে nullity হচ্ছে p, ঠিক যেটা দেখানোর দরকার ছিল। **[Q.E.D]**

THEOREM

If $\vec{v}_1,...,\vec{v}_k$ are generalised eigenvectors of a matrix for distinct eigenvalues, then $\{\vec{v}_1,...,\vec{v}_k\}$ must be linearly independent.

Proof:

> (Induction on k):
> <u>Basis:</u> For $k=1$, we have only a single non-zero vector, and so linearly independent.
> <u>Hypothesis:</u> We assume the result up to $k-1$.
> <u>Step:</u> Shall show for k.

আমরা এবার এমন $\vec{v}_1,...,\vec{v}_k$ নেব, যারা সকলেই A-র একেকটা generalised eigenvector, আলাদা আলাদা generalised eigenvalue-র জন্য।

> Let $\vec{v}_1, ..., \vec{v}_k$ be generalised eigenvectors of $A_{n\times n}$ corresponding to distinct eigenvalues.

এদেরকে linearly independent দেখাতে হবে। তার জন্য সেই চিরাচরিত কায়দা, ওদের যা খুশি একটা linear combination নিয়ে শূন্যর সাথে সমান করে দেব, এবং দেখাব যে, এর ফলে প্রতিটা coefficient শূন্য হয়ে যেতে বাধ্য।

> Let
> $$\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k = \vec{0} \qquad (*)$$
> for some scalars $\alpha_1, ..., \alpha_k$.
> Let the generalised eigenvalue for $\vec{v}_i$ be λ_i for $i = 1, ..., k$.
> Then $(A - \lambda_1 I)^n\vec{v}_1 = \vec{0}$.

এটা পেয়েছি generalised eigenvector-এর সংজ্ঞা থেকে।

> Then $(A - \lambda_1 I)^n(\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k) = (A - \lambda_1 I)^n\vec{0} = \vec{0}$.
> Hence
> $$\alpha_2\vec{(}A - \lambda_1 I)^n\vec{v}_2 + \cdots + \alpha_k(A - \lambda_1 I)^n\vec{v}_k = \vec{0}. \qquad (**)$$
> But for $i = 2, ..., k$, the vector $(A-\lambda_1 I)^n\vec{v}_i$ belongs to the same generalised eigenspace as $\vec{v}_i$.

কারণ আমরা আগেই বলেছি যে, কোনো generalised eigenspace-এর কোনো vector উপরে A-র কোনো polynomial লাগালেও সেটা ওই subspace-এই থেকে যায়।

> ⟦ Because:
> If the generalised eigenvalue for $\vec{v}_i$ is λ_i, then
> $$(A-\lambda_i I)^n(A-\lambda_1 I)\vec{v}_i = (A-\lambda_1 I)^n(A-\lambda_i I)\vec{v}_i = (A-\lambda_1 I)\vec{0} = \vec{0}.$$
> ⟧
> Also since $\vec{v}_i$'s correspond to distinct generalised eigenvalues, hence $\forall i \in \{2, 3, ..., k\} \quad (A - \lambda_1 I)\vec{v}_i \neq \vec{0}$.

কারণ, যদি কোনো $i = 2, ..., k$-এর জন্য $(A - \lambda_1 I)\vec{v}_i = \vec{0}$ হত, তবে $(A - \lambda_1 I)^n\vec{v}_i = \vec{0}$ হতেও বাধ্য। মানে $\vec{v}_i$-এর জন্য generalised eigenvalue হয়ে যাচ্ছে λ_1. কিন্তু বলা আছে যে, $\vec{v}_i$-এর জন্য generalised eigenvalue হল $\lambda_i \neq \lambda_1$.

> Hence $(A - \lambda_1 I)^n\vec{v}_2, ..., (A - \lambda_1 I)^n\vec{v}_k$ are also generalised eigenvectors corresponding to distinct eigenvalues, $\lambda_2, ..., \lambda_k$.
> So, applying induction hypothesis on (**), we have, $\alpha_2 = \alpha_k = 0$.

Hence, from (*), $\alpha_1 \vec{v}_1 = \vec{0}$, and so $\alpha_1 = 0$, because $\vec{v}_1 \neq \vec{0}$.

[Q.E.D]

এই কয়টা জিনিস মেলালেই আমরা primary decomposition theorem পেয়ে যাব--

Primary decomposition theorem

Let $A_{n\times n}$ be a matrix with distinct eigenvalues $\lambda_1, ..., \lambda_k \in \mathbb{C}$ with generalised eigenspaces $E_1, E_2, ..., E_k$, respectively. Then

$$\mathbb{C}^n = E_1 \oplus E_2 \oplus \cdots \oplus E_k.$$

এই $\oplus$-এর ব্যাপারটা মনে আছে তো? এখানে $\mathbb{C}^n = E_1 \oplus E_2 \oplus \cdots \oplus E_k$ বলা মানে দুটো জিনিস বলা--

- এক, $E_1 + \cdots + E_k = \mathbb{C}^n$
- আর দুই, যদি প্রতিটা E_i থেকে একটা করে $\vec{x}_i \neq 0$ নাও, তবে $\{\vec{x}_1, ..., \vec{x}_k\}$ হবে linearly independent.

Proof:

Clearly $E_1 + \cdots + E_k \leq \mathbb{C}^n$.

Also, since generalised eigenvectors corresponding to distinct eigenvalues are linearly independent, hence $E_1 + \cdots + E_k = E_1 \oplus \cdots \oplus E_k$

তার মানে পেলাম $E_1 \oplus E_2 \oplus \cdots \oplus E_k \leq \mathbb{C}^n$. এবার "$\leq$"-এর জায়গায় "=" দেখাতে হবে। তার জন্য dimension-এর কায়দা ব্যবহার করব। দেখাব যে, $\dim(E_1 \oplus E_2 \oplus \cdots \oplus E_k) = n$. একটা subspace-এর dimension যদি পুরো space-এর dimension-এর সমান হয়, তবে subspace-টা নিজেই পুরো space-টার সমান হতে বাধ্য।

Now $\dim(E_i) =$ algebraic multiplicity of λ_i.

Since the algebraic multiplicities add up to n,

So $\dim(E_1 \oplus \cdots \oplus E_k) = \dim(E_1) + \cdots + \dim(E_k) = n$.

Thus, $E_1 \oplus \cdots \oplus E_k = \mathbb{C}^n$, as required.

[Q.E.D]

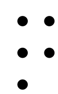
অনেকগুলো theorem শেখা হল। এবার বলি এগুলো দিয়ে আমাদের কী লাভ হবে। সেই যে প্রথমেই বলেছিলাম যে, distinct eigenvalue-গুলো নিয়ে আলাদা করে কাজ করা যায়, এই theorem-টার জোরেই সেটা বলেছিলাম। আমরা একটু আগে যখন ছকের মধ্যে কোণাকুণিভাবে নড়ে একটা basis বানিয়েছিলাম, সেই কায়দাটা এখানে প্রতিটা distinct eigenvalue-র উপরে আলাদা করে করলে E_i-গুলোর একটা করে basis বেরোবে। যেহেতু $E_1 \oplus \cdots \oplus E_k = \mathbb{C}^n$, তাই ওই basis-গুলোকে পাশাপাশি বসিয়ে দিলেই পুরো $\mathbb{C}^n$-এর জন্য উপযুক্ত একটা basis পেয়ে যাবে। সেই basis থেকে যে P পাবে, তার জন্যই $AP = JP$ হবে, যেখানে J হল A-র একটা JCF.

Exercise 50: নীচের বাক্যটা ঠিক নাকি ভুল?

"কোনো eigenvalue-র algebraic multiplicity আর তার generalised eigenspace-এর dimension সর্বদা সমান হয়।"

■

Answers

1. $k = 3$. **2.** (B). **3.** অসম্ভব, কারণ যাই A নাও না কেন, $m(A) = I$ হয়ে যাবে। **4.** না, যেমন $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. বস্তুতঃ, এরকম বিস্তর counter-example অতি সহজেই বানানো যায়--যেকোনো একটা triangular matrix নাও, যার diagonal বরাবর সব শূন্য, কিন্তু কোথাও অন্ততঃ একটা nonzero off-diagonal entry আছে।
5. **(1)** $m(x) = (x-1)(x-2)(x-3)$ **(2)** $m(x) = (x-1)^2(x-2)(x-3)$ বা $(x-1)(x-2)(x-3)$ **(3)** $m(x) = (x^2+1)(x^2+2)$ **(4)** $m(x) = x^3$ বা x^2 বা x.
6. $\begin{bmatrix} 7+2i & 1 & 0 \\ 0 & 7+2i & 1 \\ 0 & 0 & 7+2i \end{bmatrix}$. **7.** $(a-b)I_k$. **8.** $m = k$ আর $\lambda = 6 - 2 = 4$. **9.** না, দুজনেই $\begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & ? & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ চেহারার, খালি ওই ?-এর জায়গায় $J_4(3)$-এর ক্ষেত্রে 1 আছে, আর অন্যটার ক্ষেত্রে 0.
10. $\begin{bmatrix} 3 & 1 & & & & & \\ & 3 & & & & & \\ & & 3 & 1 & & & \\ & & & 3 & & & \\ & & & & 2 & 1 & \\ & & & & & 2 & 1 \\ & & & & & & 2 \end{bmatrix}$ **11.** হ্যাঁ। Jordan block-গুলো হল $J_1(1)$, $J_1(2)$, $J_1(3)$ এবং $J_1(4)$.
12. $\operatorname{diag}\big(J_3(5), J_2(5), J_2(i), J_1(i), J_1(3), J_1(3)\big)$. এখানে 5-এর জন্য বিন্দুগুলো হল  , i-এর বেলায় এবং 3-এর ক্ষেত্রে  . **13.** $\operatorname{diag}(1, 1, 2, 2)$. **14.** $J_1(2)$ আছে সাতটা, আর $J_2(2)$ আছে একটা।
15. $J_1(0)$ দুটো, $J_2(0)$ একটা, $J_1(i)$ দুটো, $J_4(-2)$ একটা। **16.** $\operatorname{diag}(1, 3, 6)$. **17.** Eigenvalue-গুলো হল 3 এবং 1. এদের algebraic multiplicity যথাক্রমে $5 \times 1 + 2 \times 3 = 11$ এবং $2 \times 2 = 4$.
18. Eigenvalue-গুলো হল i এবং 5. এদের algebraic multiplicity হল যথাক্রমে $2 + 3 = 5$ এবং $2 \times 3 = 6$. সুতরাং characteristic polynomial হচ্ছে $p(\lambda) = (\lambda - i)^5(\lambda - 5)^6$. **19.** $(4^3)^2 \times ((2i)^5)^3 = -2^{27}i$.
20. Eigenvalue-গুলো হল i, 2 এবং $-i$. এদের geometric multiplicity যথাক্রমে $5 + 2 = 7$, 3 এবং 10.
21. এখানে কোনো $J_k(0)$ চেহারার Jordan block নেই। তাই nullity হল 0. মোট column-সংখ্যা হল $3 + 1 + 1 + 3 + 3 = 11$. সুতরাং rank হল $11 - 0 = 11$. **22.** $m(\lambda) = (\lambda - i)^2(\lambda + i)^2(\lambda - 5)^3\lambda^4$.
24. হ্যাঁ, কারণ $m \leq n$ হলে সর্বদাই n-খানা জিনিসকে m-খানা গুচ্ছে ভাগ করা যাবে, যেখানে প্রতিটা গুচ্ছেই কিছু না কিছু আছে। প্রতিটা গুচ্ছ একটা করে Jordan block-এর জন্ম দেবে। এটা যদি বুঝতে অসুবিধা হয়, তবে একটা উদাহরণ নাও--ধরো $n = 6$ আর $m = 4$ নিলাম। এবার n-খানা জিনিসকে এক লাইনে রাখো। তাদের মাঝে মাঝে যে $6 - 1 = 5$-খানা ফাঁক আছে, তার থেকে যেকোনো $4 - 1 = 3$-টে দিয়ে একটা করে লাইন টানো--
•|• •|•|• •
তবেই 4-খানা গুচ্ছ হয়ে যাবে। **25.** অসম্ভব, ব্লক মোটে দুটো, তাদের সাইজ বড়জোর 4 করে। সুতরাং 3-এর সংখ্যা হতে পারে খুব বেশী হলে $2 \times 4 = 8$. **26.** JCF-এর diagonal বরাবর λ_0 আছে মোট a-খানা। সেগুলো g-খানা Jordan

block-এ বিভক্ত। সবচেয়ে বড় Jordan block-টার সাইজ যদি m বলি, তবে g-খানা block মিলিয়ে বড়জোর মোট mg-টা λ_0 থাকতে পারে। তার মানে $a \leq mg$. সুতরাং $\frac{a}{g} \leq m$. যেহেতু m একটা integer, তাই $k = \lceil \frac{a}{g} \rceil \leq m$ হবে। এদিকে আমরা জানি যে, $(\lambda - \lambda_0)^m$ হল minimal polynomial-এর একটা factor. সুতরাং $(\lambda - \lambda_0)^k$ দিয়ে minimal polynomial-টা ভাগ যেতে বাধ্য। **27.** এখানে $k > a - g + 1$ হতে পারে না, কারণ নইলে $a - k < g - 1$ হত। সুতরাং সবচেয়ে বড় ব্লকটা বাদ দিলে পড়ে থাকে $g - 1$-এর থেকে কমসংখ্যক λ_0. এদিকে এখনও বানানো বাকি আছে $g - 1$-খানা ব্লক। সেটা তো অসম্ভব, কারণ প্রত্যেকটা ব্লকে তো অন্ততঃ একটা করে λ_0 দিতে হবে! এবার দেখাতে হবে যে, $\left\{\lceil \frac{a}{g} \rceil, ..., a - g + 1\right\}$-এর ভিতরের সবগুলো value-ই k নিতে পারে। এরকম যা খুশি একটা k নেওয়া যাক। আমরা i-th গুচ্ছের সাইজ নেব--

$$n_i = \begin{cases} k & \text{if } i = 1 \\ [(i-1)s] - [(i-2)s] & \text{if } i = 2, ..., g. \end{cases},$$

যেখানে $s = \frac{a-k}{g-1}$. দেখাতে হবে $\forall i \quad 1 \leq n_i \leq k$. যদি $i = 1$ হয়, তবে কোনো সমস্যা নেই। যদি $i \geq 2$ হয়, তবে আমরা $[x]$-এর একটা ধর্ম ব্যবহার করব, যদি $x, y \in \mathbb{R}$ এবং $p, q \in \mathbb{Z}$ হয়, তবে $p \leq x - y \leq q$ হলে $p \leq [x] - [y] \leq q$ হতে বাধ্য (কেন?)। যেহেতু $(i-1)s - (i-2)s = s \in [1, k]$, তাই $n_i \in [1, k]$ হতেও বাধ্য। **28.** JCF-টায় চারটে $J_1(1)$, দুটো $J_1(i)$, দুটো $J_1(-i)$, একটা $J_2(2)$ এবং একটা $J_1(2)$ থাকবে, এবং অন্য কিছু থাকবে না। এদেরকে নানাভাবে সাজিয়ে মোট $\frac{10!}{4!2!2!}$-খানা JCF বানানো যায়। **29.** $R_1(i) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ এবং $R_1(-i) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ **30.** মূল ধাপগুলো খালি দিচ্ছি--প্রথমে eigenvalue-গুলো বার করতে হবে, 4 আর 2. এদের algebraic multiplicity হল যথাক্রমে 1 আর 2. সুতরাং 4-এর geometric multiplicity-ও 1 হবে। এবার 2-এর geometric multiplicity বার করতে হবে। তার জন্য $A - 2I$-এর দিকে তাকাও, $\begin{bmatrix} 2 & -1 & 1 \\ 4 & -2 & 2 \\ 2 & -1 & 1 \end{bmatrix}$. দেখাই যাচ্ছে যে, সবগুলো column-ই একই জিনিসের multiple. সুতরাং rank হল 1, তাই nullity হল $3 - 1 = 2$. সুতরাং অন্য eigenvalue-টাও regular. তাই matrix-টা diagonalisable. তাই JCF পেয়ে যাবে eigenvalue-গুলো দিয়ে diagonal matrix বানালেই, যেমন $\text{diag}(4, 2, 2)$ হবে একটা JCF. **31.** $\begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$. **32.** $[-a]$.

33. $x^4 - 3x^3 - 2x^2 - 2x - 1$. **34.** $\begin{bmatrix} 0 & -1 & & & \\ 1 & 0 & & & \\ & & 0 & 0 & 1 \\ & & 1 & 0 & -1 \\ & & 0 & 1 & 1 \end{bmatrix}$. ফাঁকা জায়গাগুলোতে শূন্য আছে।

35. $p(\lambda) = (\lambda^2 - \lambda - 2)^2$. **36.** $\begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$. **37.** $\text{diag}(C_{p_1(x)p_2(x)}, C_{p_1(x)p_2(x)})$.

38. $\text{diag}(C_{p(x)}, C_{p(x)}, C_{p(x)}, C_{p(x)^2}))$ বা $\text{diag}(C_{p(x)}, C_{p(x)^2}, C_{p(x)^2}))$. **39.** $C_{p(x)}$-এর মধ্যে একটা I_{n-1} আছেই, তাই $r(C_{p(x)}) \geq n - 1$. যদি $p(0) = 0$ হয়, প্রথম row-তে সবাই শূন্য। তাই rank-টা $n - 1$-এর চেয়ে বেশী হতে পারছে না। যদি $p(0) \neq 0$ হয়, তবে প্রথম row-এর শেষ entry-টা খালি nonzero, ফলে ওই row-টা বাকি row-গুলোর span-এর বাইরে পড়ছে। তাই rank-টা n হয়ে যাচ্ছে। **40.** ঠিক ওই ব্লকগুলোর প্রথম row-গুলোতে সবাই শূন্য হবে। সেই জন্য nullity অন্ততঃ k হবে। বাকি row-গুলোর মধ্যে আলাদা আলাদা জায়গায় identity matrix থাকায় ওই সব row হচ্ছে linearly independent. তাই nullity-টা k-র বেশী হতে পারছে না। **41.** যেহেতু $p_t(x)|p_{t+1}(x)| \cdots |p_k(x)$, তাই $p_t(0) = 0$ হওয়ায় $\forall i = t+1, ..., k \quad p_i(0) = 0$ হবে। সুতরাং t-th থেকে শুরু করে k-th পর্যন্ত মোট $k - t + 1$-খানা ব্লক nullity-তে 1 সরবরাহ করছে। গুছিয়ে লেখার সময়ে এটা ব্যবহার করলে সুবিধা হবে যে, block diagonal matrix-এর nullity হয় block-গুলোর nullity-র যোগফল।

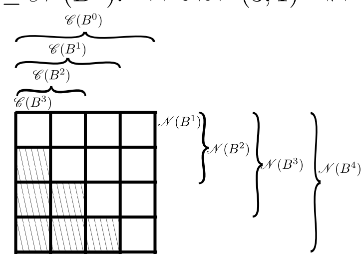
$\text{nullity}(\text{diag}(C_{p_1(x)},...,C_{p_k(x)})) = \sum \text{nullity}(C_{p_i(x)})$.

Now $\text{nullity}(C_{p(x)}) = \begin{cases} 1 & \text{if } p(0)=0 \\ 0 & \text{otherwise} \end{cases}$.

So $\text{nullity}(C_{p_i(x)}) = \begin{cases} 1 & \text{if } i=t,...,k \\ 0 & \text{otherwise} \end{cases}$.

Hence $\sum \text{nullity}(C_{p_i(x)}) = k-t+1$, as required.

42. এখানে $A\begin{bmatrix} \vec{z}_1 & \vec{z}_2 \end{bmatrix} = \begin{bmatrix} \vec{z}_1 & \vec{z}_2 \end{bmatrix}\begin{bmatrix} w & 1 \\ 0 & w \end{bmatrix}$. দুপাশের প্রথম column দুটোকে সমান করলে আসে $A\vec{z}_1 = w\vec{z}_1$. আগের উদাহরণটার মত এগোলেই real Jordan block-এর প্রথম 2×2 ব্লকটা এসে যাবে। এবার দ্বিতীয় column দুটোকে সমান করে পাচ্ছি $A\vec{z}_2 = \vec{z}_1 + w\vec{z}_2$. এর real আর imaginary অংশ আলাদা করলেই দ্বিতীয় ব্লকটা পেয়ে যাবে। **43.** $\mathscr{C}(B)$-র যেকোনো সদস্যই দেখতে $B\vec{x}$-এর মত। এবার $B^2(B\vec{x}) = B^3\vec{x} = O\vec{x} = \vec{0}$. তাই $\mathscr{C}(B) \subseteq \mathscr{N}(B^2)$. এর থেকে $(3,1)$ আর $(3,2)$ ঘর দুটোকে ফাঁকা বলা যায়।

44. 

45. একেবারে উপরের বাঁদিকের ঘরে, মানে $(1,1)$ ঘরে।

46. যেভাবে $\vec{y}$ বানালাম, তাতেই $\vec{y} \in \mathscr{C}(B^{k-j-1})$ হয়ে গেছে। আবার $\vec{x} \notin \mathscr{C}(B^{k-j+1})$ হওয়ায় $\vec{y} \notin \mathscr{C}(B^{k-j})$. সুতরাং $\vec{y}$ আছে $i+1$ নম্বর row-এর কোথাও। আবার $B^i\vec{x} = \vec{0}$, মানে $B^{i+1}\vec{y} = \vec{0}$. তাই $\vec{y} \in \mathscr{N}(B^{i+1})$. কিন্তু $\vec{y} \notin \mathscr{N}(B^i)$, কারণ নইলে $\vec{x} \in \mathscr{N}(B^{i-1})$ হয়ে যেত। সুতরাং $\vec{y}$ আছে $j+1$ নম্বর column-এ। **48.** $B\vec{c} = \vec{0}$ মানে $(A-\lambda I)\vec{c} = \vec{0}$, অর্থাৎ $A\vec{c} = \lambda\vec{c}$. আবার $B\vec{h} = \vec{c}$ থেকে পাচ্ছি $A\vec{h} = \vec{c} + \lambda\vec{h}$. সুতরাং $A\begin{bmatrix} \vec{c} & \vec{h} \end{bmatrix} = \begin{bmatrix} \vec{c} & \vec{h} \end{bmatrix}\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$. **49.** $D = \text{diag}(J_3(\lambda), J_2(\lambda))$. **50.** ঠিক।

Chapter XI
Duality

DAY 71
Duality (part 1)

এই বইতে আমরা কাজ করছি finite dimensional vector space নিয়ে। কখনো-সখনো infinite dimensional vector space-দের প্রসঙ্গ আসেনি এমন নয়, কিন্তু আমাদের মূল আলোচনায় বা উদাহরণে সব vector space-এর dimension-ই finite নিয়েছি। এই অধ্যায়েও তার ব্যতিক্রম হবে না, কিন্তু এখানে আমরা যা আলোচনা করছি, সেটা finite dimensional vector space-এর ক্ষেত্রে খুব একটা নতুন কিছু দেবে না। এর আসল চমক হল infinite dimension-এর বেলায়। সেই চমকটা কী, তার একটা আভাস দেব কিছুক্ষণ পরে। আগে finite dimensional vector space-দের বেলায় ব্যাপারটা বুঝে নিই। তার জন্য প্রথমেই জানতে হবে linear functional কাকে বলে।

71.1 Linear functional

আমরা linear transformation-এর কথা শিখেছি। ধরো আমাদের field-টা হল F. তবে কোনো $T : F^n \to F^m$-কে linear transformation বলতাম, যদি $T(\vec{x}) = A\vec{x}$ লেখা যায়, যেখানে A হল $m \times n$ সাইজের constant কোনো matrix. যদি $m = 1$ হয়, তবে সেরকম linear transformation-দেরই বলি **linear functional**.
ফর্মুলা দেখে এদের চেনা খুব সহজ। $T : F^n \to F$ একটা linear functional হওয়া মানে $T(x_1, ..., x_n)$-এর চেহারা হতে হবে এইরকম--

$$T(x_1, ..., x_n) = a_1x_1 + \cdots + a_nx_n,$$

যেখানে $a_1, ..., a_n \in F$. এই চেহারাটা মাথায় রেখে নীচের অংকটা করো।

Example 1: এদের মধ্যে কারা $\mathbb{R}^3$-এর উপরে linear functional?

(1) $f(x, y, z) = x+2y-z$ **(2)** $f(x, y, z) = 5+x+y+z$ **(3)** $f(x, y, z) = xy+z$ **(4)** $f(x, y, z) = (x, y)$ **(5)** $f(x, y, z) = x + 3z$

SOLUTION: প্রথমটা একটা linear functional, কারণ এখানে আমরা $f(x, y, z) = a_1x + a_2y + a_3z$ লিখতে পারি, যেখানে $a_1 = 1,\ a_2 = 2$ এবং $a_3 = 1$. দ্বিতীয়টার বেলায় ওই 5-টা থেকে সব মাটি করে দিয়েছে, এটা linear transformation-ই হচ্ছে না, তবে আর linear functional হয় কী করে? তৃতীয়টার বেলায় একটা xy দেখা যাচ্ছে, তাই এটাও linear transformation-ই হচ্ছে না। চতুর্থটা linear transformation বটে, কিন্তু $\mathbb{R}^3$ থেকে $\mathbb{R}^2$-তে। আমাদের দরকার ছিল $\mathbb{R}^3$ থেকে $\mathbb{R}$-এ। তাই এটাও linear functional হল না। শেষেরটার বেলায় অবশ্য কোনো সমস্যা নেই। এখানে $a_1 = 1,\ a_2 = 0$ এবং $a_3 = 3$. ■

দুটো linear functional-কে যদি যোগ করো, তবে ফের একটা linear functional-ই পাবে। যদি একটা linear functional-কে কোনো scalar দিয়ে গুণ করো, তাহলেও ফের একটা linear functional-ই আসবে। এই দুটো ব্যাপার মিলে আমরা বলতে পারি যে, যাবতীয় linear functional-রা মিলে একটা vector space হবে। এই যে যাবতীয় linear

functional-দের তৈরী vector space, এটারই পোশাকী নাম হল **dual space**. আমরা এখানে $\mathbb{R}^n$-এর উদাহরণ দিয়ে বলছি। কিন্তু dual space-এর ধারণাটা যেকোনো field-এর উপর যেকোনো vector space-এর জন্যই খাটে। সেটাই গুছিয়ে লিখেছি নীচের সংজ্ঞায়--

> DEFINITION: **Dual space**
>
> Let V be a vector space over a field F. Then the **dual space** of V is the vector space of all linear functionals over V. It is denoted by V^*.

লক্ষ্য করো, এখানে আমরা V-এর dimension-এর উপর কোনো শর্ত চাপাই নি। V-টা finite dimensional-ও হতে পারে, infinite dimensional-ও হতে পারে। এবার আমরা dual basis বলে একটা জিনিসের কথা শিখব। তার জন্য অবশ্য V-কে একটা finite dimensional vector space নেব।
যদি V-র কোনো basis দেওয়া থাকে, তবে তুমি তা থেকে V^*-এর একটা বিশেষ basis বানাতে পারবে। সেটাকে বলে **dual basis**. একটা উদাহরণ দিয়ে বলি। তার জন্য এই জিনিসটা মনে রাখতে হবে--

> যদি $\{\vec{v}_1, ..., \vec{v}_n\}$ একটা কোনো basis হয়, আর $a_1, ..., a_n \in F$ যা খুশি কিছু scalar হয় (সবগুলো আলাদা হবারও কোনো দরকার নেই), তবে ঠিক একটাই linear functional পাবে $f : V \to F$, যাতে $\forall i \quad f(\vec{v}_i) = a_i$ হয়।

এটা linear transformation-দের ধর্ম থেকেই আসে, যেটা আমরা সেই চতুর্থ অধ্যায়েই শিখেছিলাম। এই তথ্যটা কাজে লাগিয়ে এবার আমরা একটা dual basis বার করার অংক দেখি।

Example 2: ধরো $\mathbb{R}^2$-র একটা basis দিলাম, $\{\vec{v}_1, \vec{v}_2\}$, যেখানে $\vec{v}_1 = (1, 0)$ আর $\vec{v}_2 = (1, 1)$. এর dual basis বার করো।
SOLUTION: এখানে dual space-টা নিয়ে, মানে $\mathbb{R}^2$-এর উপরে যাবতীয় linear functional-এর vector space. এবার থেকে দুটো বিশেষ linear functional নেব, $\{f_1, f_2\}$, যাতে--

$$f_1(\vec{v}_1) = 1 \quad , \quad f_1(\vec{v}_2) = 0,$$
$$f_2(\vec{v}_1) = 0 \quad , \quad f_2(\vec{v}_2) = 1$$

হয়। এইটাই হল dual basis হওয়ার শর্ত। In general,

$$f_i(v_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} .$$

এই শর্ত মেনে f_1 আর f_2 যে পাওয়া যাবেই, এবং ঠিক একভাবেই পাওয়া যাবে, সেটা তো এক্ষুণিই বললাম। এরা অবশ্যই V^*-এর একটা basis হবে।
এবার হাতেকলমে বার করা যাক। $\mathbb{R}^2$-র উপরে যেকোনো linear functional-ই হল $ax + by$-জাতীয়। ধরো $f_1(x, y) = ax + by$ এবং $f_2(x, y) = cx + dy$. তবে আমরা চাইছি চারটে শর্ত, দুটো a, b-কে নিয়ে--

$$\begin{aligned} f_1(\vec{v}_1) = a \times 1 + b \times 0 &= 1, \\ f_1(\vec{v}_2) = a \times 1 + b \times 1 &= 0, \end{aligned}$$

আর দুটো c, d-কে নিয়ে--

$$\begin{aligned} f_2(\vec{v}_1) = c \times 1 + d \times 0 &= 0, \\ f_2(\vec{v}_2) = c \times 1 + d \times 1 &= 1. \end{aligned}$$

এদের সমাধান করলেই পাবে $a = 1,\ b = -1,\ c = 0$ আর $d = 1$. তার মানে $f_1(x) = ax + by = x - y$ এবং $f_2(x) = cx + dy = y$. ■

এবার তবে dual basis-এর সংজ্ঞাটা গুছিয়ে লিখে ফেলি।

DEFINITION: Dual basis

Let V be a finite dimensional vector space. Let V^* be the dual space of V. Let $\{\vec{v}_1, ..., \vec{v}_n\}$ be any basis of V. Then $\{f_1, ..., f_n\} \subseteq V^*$ is called the dual basis of $\{\vec{v}_1, ..., \vec{v}_n\}$ if

$$f_i(\vec{v}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

এই সংজ্ঞাটার বিষয়ে কয়েকটা কথা বলা দরকার--

- এক, এখানে কোথাও বললাম না যে $\{f_1, ..., f_n\}$-কে V^*-এর basis হতে হবে। কিন্তু সংজ্ঞার শর্ত মেনে যদি $\{f_1, ..., f_n\}$ নাও, তবে সেটা আপনা থেকেই হবে।
- দুই, এরকম $\{f_1, ..., f_n\}$ কি পাওয়া যাবেই? সে গ্যারান্টিটা সংজ্ঞার মধ্যে দেওয়া নেই। আসলে কিন্তু সে গ্যারান্টিটা দেওয়া যায়। তাই যেকোনো (finite) basis-এর জন্য dual basis থাকবেই।
- তিন, সংজ্ঞার শর্ত মেনে কি একাধিকভাবে $\{f_1, ..., f_n\}$ নেওয়া যেতে পারে? উত্তর হল--না, ওই শর্ত মানলে $\{f_1, ..., f_n\}$ অবশ্যই unique হবে।

এই যে তিনটে জিনিস বললাম, এগুলো প্রমাণ করা নিয়েই নীচের অংকগুলো।

Exercise 1: দেখাও যে, সংজ্ঞার শর্ত মেনে $\{f_1, ..., f_n\} \subseteq V^*$ ঠিক একভাবেই নেওয়া সম্ভব। ■

Exercise 2: দেখাও যে, $\{f_1, ..., f_n\} \subseteq V^*$ যদি সংজ্ঞার মত করে নেওয়া যায়, তবে তারা মিলে V^*-এর একটা basis হবে।
HINT: Linearly independent দেখানোর জন্য $a_1 f_1 + \cdots + a_n f_n = 0$ নিয়ে শুরু করো। এখানে ডানদিকের 0-টা মানে সেই linear functional, যেটা সর্বদাই শূন্য। দেখো তো, এ থেকে $a_1 = \cdots = a_n = 0$ দেখাতে পারো কিনা। Spanning দেখানোর জন্য যা খুশি একটা $f \in V^*$ নাও। এবার $a_i = f(\vec{v}_i)$ নিয়ে দেখাও যে $a_1 f_1 + \cdots + a_n f_n = f$ হচ্ছে। ■

Exercise 3: Choose the correct answer and justify (1 mark for right answer and 1 mark for justification). If $B = \{v_1, v_2, v_3\}$ is an ordered basis for $\mathbb{C}^3$ defined by $v_1 = (1, 0, -1)$, $v_2 = (1, 1, 1)$, $v_3 = (2, 2, 0)$, then the dual basis $\{f_1, f_2, f_3\}$ of B is given by

(A) $f_1(x, y, z) = x + y;\ f_2(x, y, z) = x + y + z;\ f_3(x, y, z) = \frac{x-y+z}{2}$

(B) $f_1(x, y, z) = x - y;\ f_2(x, y, z) = x - y + z;\ f_3(x, y, z) = -\frac{1}{2}x + y - \frac{1}{2}z$

(C) $f_1(x, y, z) = x + y;\ f_2(x, y, z) = -\frac{1}{2}x + y - \frac{1}{2}z;\ f_3(x, y, z) = x - y - z$

(D) None of the above.

(2020.CC12.1j)
HINT: প্রতি ক্ষেত্রে তোমাকে দেখতে হবে $f_i(\vec{v}_i) = 1$ এবং $i \neq j$ হলে $f_i(\vec{v}_j) = 0$ হচ্ছে কিনা। যেমন ধরো (A)-র বেলায় $f_2(\vec{v}_2) = 1 + 1 + 1 = 3$. সুতরাং (A)-টা উত্তর হতে পারে না। ■

Exercise 4: Let $V = \mathbb{R}^3$ be the vector space over $\mathbb{R}$ and V^* be its dual space. Let $f_1, f_2, f_3 \in V^*$ be such that $f_1(x,y,z) = x - 2y,\ f_2(x,y,z) = x + y + z,\ f_3(x,y,z) = y - 3z$. Prove that $\{f_1, f_2, f_3\}$ is a basis for V^* and find a basis for V for which it is the dual basis. [5]
(2021.CC12.3d)
HINT:
প্রথমে basis দেখাব। তার জন্য linearly independent আর spanning দেখাতে হবে। দুটোই এক সঙ্গে দেখানো যাবে--

> Let us write $f = af_1 + bf_2 + cf_3$ for some scalars a, b, c.
> Then $f(x,y,z) = \cdots = (a+b)x + (-2a+b+c)y + (b-3c)z$.
> Setting $\vec{e}_1, \vec{e}_2, \vec{e}_3$ one by one in place of (x,y,z), we have
>
> $$\begin{aligned} a + b &= f(1,0,0), \\ -2a + b + c &= f(0,1,0), \\ b - 3c &= f(0,0,1), \end{aligned}$$
>
> or
>
> $$\underbrace{\begin{bmatrix} 1 & 1 & 0 \\ -2 & 1 & 1 \\ 0 & 1 & -3 \end{bmatrix}}_{A} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} f(1,0,0) \\ f(0,1,0) \\ f(0,0,1) \end{bmatrix}.$$
>
> Here $|A| \neq 0$.

এখানে তুমি determinant-টা বার করে দেখাবে।

> Hence, if we take $f \equiv 0$, then $a = b = c = 0$. So $\{f_1, f_2, f_3\}$ is linearly independent.
> Also for any $f \in V^*$ the system is consistent, and so $\{f_1, f_2, f_3\}$ is spanning.

যেহেতু A-টা nonsingular, তাই এই system-টার ডানদিকে যাই নাও না কেন, system-টা consistent হবেই। মানে যেকোনো f-কেই f_1, f_2 আর f_3-র linear combination হিসেবে লেখা যাবে। সেইজন্যই spanning হচ্ছে।

> Thus $\{f_1, f_2, f_3\}$ is a basis for V^*, as required.

এবার দ্বিতীয় অংশ--

> Let the required basis be $\{(a_1, a_2, a_3),\ (b_1, b_2, b_3),\ (c_1, c_2, c_3)\}$.

Then

$$f_1(a_1,a_2,a_3)=1 \quad \text{or} \quad a_1-2a_2=1,$$
$$f_2(a_1,a_2,a_3)=0 \quad \text{or} \quad a_1+a_2+a_3=0,$$
$$f_3(a_1,a_2,a_3)=0 \quad \text{or} \quad a_2-3a_3=0.$$

Solving, we get $a_1=-4a_3$ and $a_2=3a_3$. Since $a_1-2a_2=1$, we have $-10a_3=1$, or $a_3=-\frac{1}{10}$. Hence $a_1=\frac{2}{5}$ and $a_2=-\frac{3}{10}$.

Again,

$$f_1(b_1,b_2,b_3)=0 \quad \text{or} \quad b_1-2b_2=0,$$
$$f_2(b_1,b_2,b_3)=1 \quad \text{or} \quad b_1+b_2+b_3=1,$$
$$f_3(b_1,b_2,b_3)=0 \quad \text{or} \quad b_2-3b_3=0.$$

Solving, we get $b_1=6b_3$ and $b_2=3b_3$. Since $b_1+b_2+b_3=1$, we have $10b_3=1$, or $b_3=\frac{1}{10}$. Hence $b_1=\frac{3}{5}$ and $b_2=\frac{3}{10}$.

একইভাবে c_1,c_2,c_3-ও বার করা যাবে। সেটা তোমার জন্য রেখে দিলাম। করার পরে উত্তরের সাথে মিলিয়ে নিও। ■

Exercise 5: Consider a basis $\{(a_1,a_2,a_3), (b_1,b_2,b_3), (c_1,c_2,c_3)\}$ of $\mathbb{R}^3$. Let the dual basis be $\{f_1,f_2,f_3\}$, where

$$f_1(x,y,z)=p_1x+p_2y+p_3z, \quad f_2(x,y,z)=q_1x+q_2y+q_3z, \quad f_3(x,y,z)=r_1x+r_2y+r_3z.$$

Show that $\{(p_1,p_2,p_3), (q_1,q_2,q_3), (r_1,r_2,r_3)\}$ is a basis of $\mathbb{R}^3$, and its dual basis is $\{g_1,g_2,g_3\}$, where

$$g_1(x,y,z)=a_1x+a_2y+a_3z, \quad g_2(x,y,z)=b_1x+b_2y+b_3z, \quad g_3(x,y,z)=c_1x+c_2y+c_3z.$$

HINT: সরাসরি সংজ্ঞা থেকে পরীক্ষা করে গেলেই হবে। ■

71.2 Matrix-এর inverse দিয়ে dual basis বার করা

আমরা যেভাবে গুচ্ছ-গুচ্ছ equation ব্যবহার করে dual basis বার করছিলাম, সেভাবে একটু বেশী dimension-এর বেলায় করা খুবই কঠিন। কিন্তু matrix-এর inverse ব্যবহার করলে অতি দ্রুত বড় বড় dimension-এর বেলাতেও dual basis বার করে ফেলা যাবে। এ কায়দাটা অবশ্য সিলেবাসে নেই। সুতরাং পরীক্ষা পাশের জন্য এটা না জানলেও চলবে।
আমরা প্রথমে একটা ছোটো উদাহরণ দিয়েই কায়দাটা বোঝাই। একবার বুঝে গেলে বড় উদাহরণে লাগাতে কোনো অসুবিধা হবে না।

Example 3: আমরা আগের একটা অংকে $\mathbb{R}^2$-র এই basis-টা দিয়েছিলাম, $\{\vec{v}_1,\vec{v}_2\}$, যেখানে $\vec{v}_1=(1,0)$ আর $\vec{v}_2=(1,1)$. এর dual basis পেয়েছিলাম $f_1(x,y)=x-y$ আর $f_2(x,y)=y$. এবার matrix-এর inverse দিয়ে বার করে দেখো একই জিনিস পাও কিনা।

SOLUTION: Matrix-এর inverse দিয়ে dual basis বার করার প্রথম ধাপ হল মূল basis-টাকে row হিসেবে একটার নীচে একটা লিখে একটা matrix বানিয়ে নেওয়া--

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

এবার এটার inverse বার করো--

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

এই inverse-টার column-গুলোকে vector হিসেবে আলাদা করে নাও, $\vec{w}_1 = (1,1)$ এবং $\vec{w}_2 = (0,1)$. এদেরকে ব্যবহার করে dual space-এর linear functional-গুলো বানাও, যেমন $\vec{w}_1$ থেকে $f_1(x,y) = \mathbf{1}\times x + \mathbf{1}\times y = x+y$. একইভাবে $\vec{w}_2$ থেকে $f_2(x,y) = \mathbf{0}\times x + \mathbf{1}\times y = y$. ঠিক এটাই আমরা আগেও পেয়েছিলাম। ■

Exercise 6: $\mathbb{R}^3$-এর এই basis-টা দিচ্ছি--

$$\{(1,2,3),(4,5,6),(7,8,0)\}.$$

এর dual basis-টা বার করো। বলে দেওয়া আছে যে,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix}^{-1} = \frac{1}{9}\begin{bmatrix} -16 & 8 & -1 \\ 14 & -7 & 2 \\ -1 & 2 & -1 \end{bmatrix}.$$

■

71.3 V-এর সঙ্গে V^*-এর তুলনা

V আর V^* দুজনেই একই field-এর উপরে দুটো vector space. যদি V-টা finite dimensional হয়, তবে V^* হয় V-এর সঙ্গে isomorphic. একটা উদাহরণ করলেই ব্যাপারটা স্পষ্ট হবে।

Example 4: মনে করো $V = \mathbb{R}^2$. এর প্রতিটা সদস্যকেই আমরা (a,b) বলে ভাবতে পারি। অর্থাৎ একটা vector-কে বোঝাতে দুটো scalar লাগছে। এখানে V^* হল $\mathbb{R}^2$ থেকে $\mathbb{R}$-এ linear transformation-রা, যাদের চেহারা হল $f(x,y) = ax+by$-এর মত। তার মানে V^*-এর একটা সদস্যকে বোঝাতেও ঠিক দুটো scalar-ই লাগছে, a আর b. যদি V-এর মধ্যে দুটো vector-কে যোগ করি, সেটা এইরকম--

$$(a,b)+(c,d) = (a+c, b+d).$$

তেমনি যদি V^*-এর মধ্যে দুটো vector নিই, $f(x,y) = ax+by$ এবং $g(x,y) = cx+dy$, তবে

$$f(x,y)+g(x,y) = (a+c)x + (b+d)y.$$

অর্থাৎ দুইক্ষেত্রেই প্রথম দুটো সংখ্যা আলাদা করে, এবং দ্বিতীয় সংখ্যাদুটো আলাদা করে যোগ হয়ে গেল। Scalar multiplication-এর বেলাতেও একই ব্যাপার--যখন V নিয়ে কাজ করছি, তখন এইরকম--

$$\alpha(a,b) = (\alpha a, \alpha b),$$

আর V^*-এর বেলায়--

$$\alpha\, f(x,y) = (\alpha a)x + (\alpha b)y.$$

তবেই বুঝতে পারছ, vector space হিসেবে V আর V^*-এর গঠন একই, মানে ওরা isomorphic. ■

> **THEOREM**
>
> Let V be a finite dimensional vector space. Let V^* be its dual space. Then V^* is isomorphic to V. We write $V^* \cong V$.

Proof: এখানে V আর V^*-কে isomorphic দেখানোর জন্য হাতেকলমে একটা isomorphism বার করে দেখানোর দরকার নেই। খালি এই তথ্যটাই যথেষ্ট যে, ওরা দুজনেই একই field-এর উপরে দুটো vector space, এবং ওদের দুজনের dimension সমান। আমরা জানি যে, F যদি কোনো field হয়, তবে তার উপরে n-dimension-এর যেকোনো vector space-ই F^n-এর সঙ্গে isomorphic হতে বাধ্য। এই কথাটাই গুছিয়ে লিখব এবার--

> We take any basis $\{\vec{v}_1, ..., \vec{v}_n\}$ of V.
>
> Then we know that V^* has a dual basis $\{f_1, ..., f_n\}$.

এইটাই হল মোক্ষম ধাপ। V-এর basis-এর যা সাইজ, dual basis-এর সাইজও তাই।

> Thus, both V and V^* are n-dimensional vector spaces over the same field, F, say.
>
> Hence, by standard result, $V \cong F^n$ and $V^* \cong F^n$.
>
> So $V \cong V^*$, as required.

[Q.E.D]

আমরা finite dimensional vector space-দের বেলায় dual space নিয়ে বড় একটা মাথা ঘামাই না। তার প্রধান কারণ হল এই theorem-টাই। V^*-এর মধ্যে গঠনগতভাবে V-এর থেকে নতুন কিছু থাকে না। Dual space-এর মূল প্রয়োগ হল infinite dimensional vector space-এর জন্য। যদিও সেই প্রয়োগটা আমাদের সিলেবাসে নেই, তাও একটু আলোচনা করছি। যাদের পরীক্ষাপাশের তাড়া, তারা এখানে কানে আঙুল দিতে পারো।
প্রথমে infinite dimensional vector space-দের বেলায় linear functional-এর কিছু উদাহরণ দেখি।

Example 5: মনে করো V হল যাবতীয় real polynomial-এর vector space. এর উপর কিছু linear functional দাও।

SOLUTION: মনে রেখো V-এর প্রতিটা element হল একটা polynomial. তার মানে যদি কোনো $f \in V^*$ নাও, তবে $f : V \to \mathbb{R}$-এর কাজ হল একটা polynomial-কে একটা real number-এ পরিণত করা। এরকম কিছু উদাহরণ হল--

- $f(v) = v(2)$, মানে v যদি একটা polynomial হয়, তবে আমরা $x = 2$-তে $v(x)$ বার করছি (যাকে ইংরাজীতে বলে evaluate করা)। যেমন $v(x) = 1 + 2x + 5x^3$, হলে $f(v) = v(2) = 45$. যদি $v_1(x), v_2(x)$ দুটো polynomial আর a_1, a_2 দুটো scalar হয়, তবে দেখাই যাচ্ছে যে,

$$f(a_1v_1 + a_2v_2) = (a_1v_1 + a_2v_2)(2) = a_1v_1(2) + a_2v_2(2).$$

  এখানে অবশ্য $x = 2$-এর আলাদা কোনো গুরুত্ব নেই। আমরা x-এর যেকোনো একটা নির্দিষ্ট value-তে evaluate করলেই একটা linear functional পেতাম। এই ধরণের linear functional-দের বলে **evaluation map.**

- $f(v) = v'(33)$, মানে $v(x)$-কে প্রথমে differentiate করে তারপর $x = 3$-তে evaluate করা হচ্ছে। যেমন $v(x) = 1 + 2x + 5x^3$, হলে $f(v) = 137$.

- $f(v) = (v$-এর মধ্যে x^4-এর coefficient$)$. যেমন $v(x) = 1 + 2x + 5x^3$, হলে $f(v) = 0$.

■

আমরা দেখেছি যে, V একটা finite dimensional vector space হলে $V \cong V^*$ হয়। কিন্তু V-টা infinite dimensional হলে সেটা কখনোই হবে না। সেটা আমরা সব infinite dimensional vector space-এর জন্য প্রমাণ করব না, খালি আগের অংকের V-টার জন্য দেখাব। সেটা নিয়েই নীচের অংকটা।

Exercise 7: V নিচ্ছি যাবতীয় real polynomial-এর vector space. ধরো $e_a : V \to \mathbb{R}$ হল a-তে evaluation map. মানে $e_a(v) = v(a)$. আমরা আগেই দেখেছি যে, $e_a \in V^*$. দেখাও যে, $\{e_a \ : \ a \in [0,1]\}$ হল V^*-এর মধ্যে linearly independent. এই set-টা uncountable. এ থেকেই প্রমাণ হয়ে যাচ্ছে যে, $V^* \ncong V$. কী করে?

HINT:

Linearly independent দেখানো মানে এটা দেখানো যে, যেকোনো finite-সংখ্যক আলাদা আলাদা (distinct) $a_1, ..., a_n \in \mathbb{R}$ নিলে $\{e_{a_1}, ..., e_{a_n}\}$ একটা linearly independent set. অর্থাৎ যদি $b_1, ..., b_n \in \mathbb{R}$ এমন হয়, যাতে $\sum_i b_i e_{a_i} \equiv 0$ হয় (মানে যাই polynomial নাও $v(x)$, যদি $\sum_i b_i e_{a_i}(v) = 0$ হয়), তবে b_i-রা প্রত্যেকেই শূন্য হতে বাধ্য। ■

71.4 করে কী লাভ?

সত্যি কথা বলতে finite dimensional vector space-দের বেলায় dual space বার করে খুব যে একটা লাভ আছে, এমন নয়। কিন্তু প্রথম যখন dual space-এর ধারণাটা এসেছিল, তখন সেটা finite dimensional vector space-দের বেলাতেই এসেছিল। আমরা এরকম দুটো প্রয়োগের কথা বলব এবার।

71.4.1 Basis দিয়ে প্রকাশ

ধরো তোমাকে একটা basis দিলাম $\{\vec{v}_1, ..., \vec{v}_n\}$ এবং একটা vector দিলাম $\vec{v}$. তবে তুমি জানো যে, $\vec{v}$-কে এইভাবে লেখা সম্ভব--

$$\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n,$$

যেখানে $a_1, ..., a_n$ হল উপযুক্ত কিছু scalar. এই a_i-গুলোর অস্তিত্ব নিয়ে সন্দেহ নেই, কিন্তু বার করার জন্য n-খানা unknown-ওয়ালা n-খানা equation-এর সমাধান করতে হবে, যেটা খুব সহজ নাও হতে পারে। এবং কোনো একটা $\vec{v}$-এর জন্য যদি সেই পরিশ্রমটাও করোও, তাও আবার নতুন কোনো $\vec{w}$ দিলে, ফের গোড়া থেকে শুরু করতে হবে। কিন্তু যদি তুমি $\{\vec{v}_1, ..., \vec{v}_n\}$-এর একটা dual basis জানো, ধরো $\{f_1, ..., f_n\}$, তবে কাজটা অনায়াসে হবে। কারণ $a_i = f_i(\vec{v})$ হবে। এমনটা কেন হবে, সেটা নিয়েই নীচের অংকটা।

Example 6: Let V be a vector space over F, $\beta = \{v_1, v_2, ..., v_n\}$ be a basis of V and $\beta^* = \{f_1, f_2, ..., f_n\}$ be the dual basis. Then show that for every $v \in V$,

$$v = f_1(v)v_1 + f_2(v)v_2 + \cdots + f_n(v)v_n.$$

[2] **(2020.CC12.3ei)**

SOLUTION:

Take any $\vec{v} \in V$. Then $\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$ for some scalars a_i's.

Since $\{f_1, ..., f_n\}$ is the dual basis, hence $\forall i, j \quad f_i(\vec{v}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$.

So

$$\begin{aligned} f_i(\vec{v}) &= f_i(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) \\ &= \sum_j a_j f_i(\vec{v}_j) \quad [\because f \text{ is linear}] \\ &= a_i. \end{aligned}$$

Hence the result.

■

এবার হাতেকলমে একটা প্রয়োগ করে দেখি।

Exercise 8: তোমাকে $\mathbb{R}^4$-এর একটা basis দিচ্ছি, $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$, যেখানে--

$$\vec{v}_1 \left(-\tfrac{1}{2}, 0, 0, \tfrac{1}{2}\right), \qquad \vec{v}_2 = \left(\tfrac{11}{20}, \tfrac{1}{10}, -\tfrac{1}{10}, -\tfrac{3}{20}\right),$$
$$\vec{v}_3 = \left(\tfrac{23}{20}, -\tfrac{7}{10}, -\tfrac{3}{10}, \tfrac{1}{20}\right), \qquad \vec{v}_4 = \left(-\tfrac{3}{4}, \tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{4}\right).$$

বলে দেওয়া আছে যে, এর dual basis হল $\{f_1, f_2, f_3, f_4\}$, যেখানে

$$\begin{aligned} f_1(x,y,z,w) &= x + y + 2z + 3w, \\ f_2(x,y,z,w) &= 2x + 3y + z + 2w, \\ f_3(x,y,z,w) &= x - y + 3z + w, \\ f_4(x,y,z,w) &= x + 4z + w. \end{aligned}$$

এই তথ্য ব্যবহার করে $(1,2,3,4)$-কে $\mathcal{B}$-এর সাপেক্ষে প্রকাশ করো। ■

71.4.2 চট্ করে inner product বার করা

এবার যে প্রয়োগটা বলব, সেটা আইনস্টাইন তাঁর relativity-র গবেষণায় ব্যবহার করেন। ব্যাপারটা বোঝার জন্য Fig 1 দেখো। এখানে দুটো vector রয়েছে, $\vec{u}$ আর $\vec{v}$. এদের Euclidean inner product কী হবে? জ্যামিতিকভাবে ভাবলে উত্তর হল--

$$\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\| \cdot \|\vec{v}\| \cdot \cos(\theta). \tag{*}$$

একই উত্তর linear algebra দিয়ে বার করতে চাইলে আমরা দুজনকেই Euclidean basis-এর সাপেক্ষে প্রকাশ করে নেব (মানে Fig 2-এর মত গ্রাফ কাগজে বসাব)। ফলে $\vec{u}$ হয়ে যাবে $(3,1)$ এবং $\vec{v}$ পরিণত হবে $(1,4)$-এ। এই দুটো tuple-এর dot product নিলেই $\vec{u}$ আর $\vec{v}$-র inner product বেরিয়ে যাবে--

$$\langle \vec{u}, \vec{v} \rangle = (3,1) \bullet (1,4) = 3 \times 1 + 1 \times 4 = 7. \tag{**}$$

যদি আমরা স্কেল-চাঁদা বসিয়ে দৈর্ঘ্য আর কোণ মেপে (*) ব্যবহার করতাম, তাহলেও একই উত্তর পেতাম। কিন্তু (**) ব্যবহার করাই সাধারণতঃ বেশী সুবিধাজনক।

Fig 1

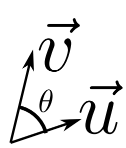
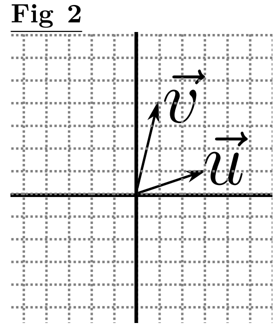


Fig 2



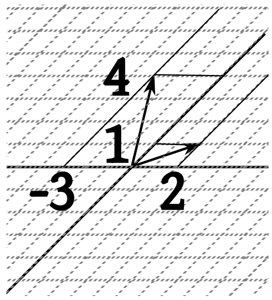


Fig 3

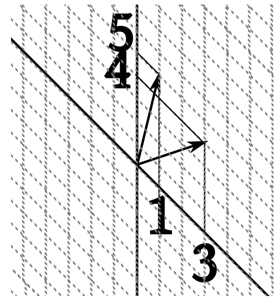


Fig 4

কিন্তু এই সুবিধার পথে একটা বাধা আছে। Fig 2-এ আমরা গ্রাফ কাগজটা এঁকেছি Euclidean basis-এর উপর ভিত্তি করে। Physics করার সময়ে সাধারণতঃ অংক কষা হয় কোনো বাস্তব সমস্যার সমাধান করার জন্য। সেখানে basis-এর নির্বাচনটা ঠিক হয়ে যায় সেই বাস্তবের প্রয়োজন থেকেই, চাইলেই আমরা Euclidean basis নিয়ে কাজ করতে পারি না। যেমন ধরো কোনো ক্ষেত্রে হয়তো দেখলে যে, basis হিসেবে $\{(1,0),(1,1)\}$ নিয়ে কাজ করতে হচ্ছে। তবে গ্রাফ কাগজটা দেখাবে Fig 3-এর মত। এখানেও $\vec{u}$ আর $\vec{v}$ একই, তাই ওদের inner product এখানেও সেই 7-ই হওয়া উচিত (কারণ দৈর্ঘ্য আর কোণ তো বদলায়নি)। কিন্তু Fig 3 থেকে দেখা যাচ্ছে যে, ওদের tuple-গুলো বদলে হয়ে গেছে $(2,1)$ এবং $(-3,4)$. (**)-এর নকলে এই tuple-দুটোর dot product নিলে কিন্তু মোটেই inner product-টা বেরোবে না--

$$\langle\vec{u},\vec{v}\rangle \neq (2,1)\bullet(-3,4) = 2\times -3 + 1\times 4 = -2.$$

এটাই দুঃখ-- আমরা dot product-ও ব্যবহার করতে চাই, আবার Euclidean basis ছাড়া অন্য কোনো basis-ও ব্যবহার করতে চাই। কিন্তু সে দুটো বিলাসিতা তো একইসঙ্গে সম্ভব বলে মনে হচ্ছে না! এইখানে আইনস্টাইন একঢিলে দুই পাখি মেরেছিলেন, দুটোকেই একই সঙ্গে সম্ভব করেছিলেন একটা কৌশল করে। তার জন্য তিনি প্রথমে basis-টার dual basis বার করে নিয়েছিলেন। Basis-টা ছিল $\{(1,0),(1,1)\}$. আমরা আগেই দেখেছি যে, এর dual basis হল $f_1(x,y)=x-y$ আর $f_2(x,y)=y$ দিয়ে তৈরী। এদেরকে যদি $ax+by$ আকারে ভাবি, তবে f_1-এর বেলায় $(a,b)=(1,-1)$ আর f_2-র বেলায় $(a,b)=(0,1)$. এই দুটো vector মিলে ফের $\mathbb{R}^2$-র একটা basis হয়। সেই অনুযায়ী গ্রাফকাগজ আঁকলে হবে Fig 4-এর মত। তার মানে আমাদের হাতে এখন দুটো basis আছে--এক, যেটা নিয়ে শুরু করেছিলাম, আর দুই, এ নতুন basis-টা। এবার আইনস্টাইন করতেন কি, $\vec{u}$-কে প্রথমটার সাপেক্ষে আর $\vec{v}$-কে দ্বিতীয়টার সাপেক্ষে লিখতেন। এর ফলে তিনি দুটো tuple পেতেন, যেমন Fig 3 থেকে $\vec{u}$-এর বেলায় tuple পাচ্ছি $(2,1)$, আর Fig 4 থেকে $\vec{v}$-এর বেলায় $(1,5)$. এই দুটো tuple-এর dot product নিলেই inner product বেরিয়ে যাবে--

$$(2,1)\bullet(1,5) = 2\times 1 + 1\times 5 = 7.$$

এই ব্যাপারটা in general যেকোনো finite dimensional vector space-এর জন্যই হবে। পুরো জিনিসটা গুছিয়ে লেখার জন্য আইনস্টাইন কিছু notation ব্যবহার করতেন। শুরুর basis-টাকে যদি $\mathcal{B}$ নাম দিই, আর নতুন basis-টাকে $\mathcal{D}$, তবে যেকোনো vector-কেই এদের যেকোনোটার সাপেক্ষে tuple-রূপে প্রকাশ করা যায়। যদি $\vec{a}$ এরকম কোনো vector হয়, তবে $\mathcal{B}$-এর সাপেক্ষে লেখার সময়ে উনি tuple-টাকে লিখতেন $(a^1,a^2,...,a^n)$. এর নাম দিয়েছিলেন $\vec{a}$-র **contravariant representation** লক্ষ্য করো, এখানে $a_1,a_2,...$ না লিখে $a^1,a^2,...$ লেখা হয়েছে। সাবধান, a^2 মানে কিন্তু মোটেই a-র square নয়। একই $\vec{a}$-কে $\mathcal{D}$-এর সাপেক্ষে প্রকাশ করলে যে tuple-টা পাওয়া যেত, তাকে লিখতেন $(a_1,a_2,...,a_n)$. এর নাম হল $\vec{a}$-র **covariant representation**. সুতরাং ওনার কায়দায়--

$$\left\langle\vec{a},\vec{b}\right\rangle = \sum_{i=1}^{n} a^i b_i.$$

এই sum-টা ওনার গবেষণায় এত আসত যে, উনি আর কষ্ট করে $\sum_{i=1}^{n}$ লিখতেন না, স্রেফ a^ib_i লিখে ছেড়ে দিতেন। গোড়াতেই কোথাও বলে রাখতেন-- "যদি কোথাও a^ib_i বা x_jy^j জাতীয় কিছু লেখা দেখো, যেখানে একই হরফ একবার

উপরতলায় একবার নীচেরতলায় রয়েছে, অমনি মনে মনে তার উপরে একটা $\sum$ বসিয়ে নেবে।" এই প্রথাটাকে এখন বলে **Einstein summation convention**.
বলাই বাহুল্য এইরকম notation ব্যবহার করায় অনভিজ্ঞ মানুষের পক্ষে তাঁর গবেষণাপত্রর মাথামুণ্ডু উদ্ধার করা বড়ই কঠিন হত। শুনেছি এই notation নাকি প্রথমদিকে সম্পাদকমশাইরাও কিছু বুঝতে পারতেন না, তাই তাঁর গবেষণাপত্র নাকচ করে ফেরত পাঠিয়ে দিয়েছিলেন!

DAY 72
Duality (part 2)

72.1 Dual of a linear transformation

আমরা একটা vector space-এর dual-এর কথা বলেছি, একটা basis-এর dual-এর কথা বলেছি। এবার একটা linear transformation-এর dual-এর কথা বলব। ধারণাটা finite dimensional বা infinite dimensional, দুই ক্ষেত্রেই একইভাবে খাটে।
ধরো $T:U\to V$ হল একটা linear transformation, যেখানে U,V হল একই field-এর উপরে যেকোনো দুটো vector space. তবে T-এর dual হল একটা বিশেষ linear transformation $T^*:V^*\to U^*$. এর সংজ্ঞাটা এইরকম--

$$T^*(f)(u)=f(T(u)) \text{ যেখানে } f\in V^*. \qquad (*)$$

অনেক সময়ে এই T^*-কে T-এর transpose-ও বলে। কেন বলে, সেটা একটু পরে বলছি। আপাততঃ এই সংজ্ঞাটাকে হজম করা যাক। আমরা বলেছি $T^*:V^*\to U^*$. তার মানে V^*-এর কোনো সদস্যের উপর T^* লাগালে U^*-এর একটা সদস্য পাওয়া যায়। V^*-এর কোন্ সদস্য থেকে U^*-এর কোন্ সদস্য আসবে সেটা বোঝার জন্য আমাদের V^*-এর কোনো একটা সদস্য নিয়ে শুরু করতে হবে, ধরো f. যেহেতু $f\in V^*$, তাই f নিজেই একটা linear functional, যেটা V-এর সদস্যদেরকে scalar-এ পরিণত করে। $T^*(f)$-কে হতে হবে U^*-এর সদস্য, মানে U-এর উপরে একটা linear functional. তার মানে $T^*(f)$-এর পেটে U-এর সদস্য ঢোকানো যাবে, এবং এর ফলে যে $T^*(f)(u)$ পাবে, সেটা হবে একটা scalar. (*) বলছে যে, সেই scalar-টাকে হতে হবে $f(T(u))$. নিশ্চয়ই গুলিয়ে যাচ্ছে? একটা উদাহরণ দেখা যাক।

Example 7: ধরো $U=\mathbb{R}^2$ এবং $V=\mathbb{R}^3$. আমাদের linear transformation-টা দিলাম এইটা, $T:U\to V$, যেখানে

$$T(x,y)=(x,y,2x+3y).$$

এর dual-টা বার করো।
SOLUTION: এর dual-টা হবে ফের আরেকটা linear transformation, যার domain হবে V^* এবং codomain হবে U^*. আমরা জানি V^*-এর সদস্যরা দেখতে সকলেই এইরকম linear functional-এর মত $f(x,y,z)=ax+by+cz$, যেখানে $a,b,c\in\mathbb{R}$. এর উপরে T^* লাগালে আসা উচিত U^*-এর কোনো সদস্য মানে $g(x,y)=px+qy$ জাতীয় একটা function. সংজ্ঞাটা (মানে (*)) বলছে--

$$T^*(f)(x,y)=f(T(x,y)).$$

এখানে $T(x,y)=(x,y,2x+3y)$. সুতরাং $f(T(x,y))=f(x,y,2x+3y)=ax+by+c(2x+3y)=(a+2c)x+(b+3c)y$. তার মানে T^*-এর প্রভাবে $f(x,y,z)=ax+by+cz$ পরিণত হচ্ছে $g(x,y)=(a+2c)x+(b+3c)y$-তে। ■

এই জায়গাতে মাথাটা ঘুরে যেতে চায়। নিজে একটা অংক করে মাথাটাকে সোজা করে নাও।

Exercise 9: যদি $T(x,y)=(x+y,x-y,2x)$ হয়, আর $f(x,y,z)=x+y+z$, তবে $T^*(f)=?$
HINT: উত্তরটা হবে $ax+by$ জাতীয়, যেখানে a,b-র জায়গায় দুটো সংখ্যা থাকবে। ■

আচ্ছা, এত কষ্ট করে এটা বার করে লাভটা কী হচ্ছে? খুব যে একটা লাভ হচ্ছে বলা যায় না। যে কারণে অনেক বইতেই এর বড় একটা আলোচনা থাকে না। তবে মূল উদ্দেশ্যটা হল matrix-এর transpose-এর ধারণাটাকে infinite dimensional vector space-এর জন্য generalise করা। সে কাজটা আমরা একভাবে আগেই করতে শিখেছি, adjoint দিয়ে। এটা আরেকটু বেশী general, কারণ এখানে কোনো inner product লাগছে না। অবশ্য আমরা (*)-এ linear transformation-এর dual-এর যে বদখত সংজ্ঞাটা দিয়েছি, তার সাথে matrix-এর transpose-এর কোনো সাদৃশ্য চোখে পড়ছে না। সেই সাদৃশ্যটা এবার বলি--

> ধরো $T : U \to V$ দেওয়া আছে, যেখানে U, V দুজনেই finite dimensional. তুমি U আর V-এর একটা করে basis নাও। তাদের সাপেক্ষে T-এর একটা matrix পাবে, ধরো A. এবার (*) অনুযায়ী T^* বার করো। সেটাও একটা linear transformation. U আর V-এর যে basis-গুলো নিয়েছিলে, তাদের dual বার করো। তারা হবে U^* আর V^*-এর basis. এই basis দুটোর সাপেক্ষে যদি T^*-এর matrix বার করো, তবে দেখবে উত্তর আসছে A'.

এই ব্যাপারটা যেকোনো field-এর জন্যই হবে। এখানেই adjoint-এর সঙ্গে dual-এর আরেকটা পার্থক্য। যদি field-টা $\mathbb{C}$ হয়, এবং তুমি ONB নিয়ে কাজ করো, তবে adjoint-টা দেবে transpose-এর conjugate. কিন্তু dual-টা সর্বদাই transpose-ই থাকবে।

একটা উদাহরণ দেখা যাক। এই উদাহরণে আমরা একটা linear transformation এবং তার dual-এর matrix বার করে দেখব যে, সত্যিই একটা আরেকটার transpose হচ্ছে।

Example 8: ধরো $U = \mathbb{R}^2$ এবং $V = \mathbb{R}^3$. আমাদের linear transformation-টা নিলাম আগের উদাহরণটার মতই, $T : U \to V$, যেখানে

$$T(x, y) = (x, y, 2x + 3y).$$

এর dual-টা আমরা পেয়েছিলাম--

$$T^*(ax + by + cz) = (a + 2c)x + (b - 3c)y.$$

এবার U আর V-এর canonical basis-এর সাপেক্ষে T-এর matrix বার করো, এবং তাদের dual basis-এর সাপেক্ষে T^*-এর। দেখো matrix দুটো পরস্পরের transpose হচ্ছে কিনা।

SOLUTION: Canonical basis-এর সাপেক্ষে T-এর matrix বার করা খুবই সহজ--

$$T(x, y) = (x, y, 2x + 3y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Canonical basis-দের dual basis বার করাও সহজই। U^*-এর বেলায় $\{x, y\}$ আর V^*-এর বেলায় $\{x, y, z\}$ (কেন?)। এদের সাপেক্ষে

$$T^*(ax + by + cz) = (a + 2c)x + (b + 3c)y$$

মানে হল (a, b, c)-কে $(a + 2c, b + 3c)$-তে নিয়ে ফেলা। যেহেতু

$$\begin{bmatrix} a + 2c \\ b + 3c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix},$$

তাই T^*-এর matrix হচ্ছে $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$, যেটা সত্যিই T-এর matrix-টার transpose. ■

এই আলোচনায় আমরা dual basis এবং matrix ব্যবহার করেছি, তাই এখানে dimension-টা finite হওয়া দরকার ছিল। কিন্তু linear transformation-এর dual-এর সংজ্ঞায় সেই প্রয়োজন নেই। সেই জন্যই linear transformation-এর dual-কে অনেক সময়ে transpose বলা হয়, কারণ ওটা transpose-এর ধারণাটারই একটা generalisation.

72.2 Annihilator

এবার আমরা একটা নতুন ধারণা শিখব, যেটা কোনো vector space এবং তার dual-এর মধ্যে একটা যোগসূত্র রচনা করবে। যেই vector space-টার মধ্যে কোনো set দেব (ধরো S), অমনি এই ধারণাটা কাজে লাগিয়ে তুমি dual-এর মধ্যে একটা set পাবে, যাকে বলব S-এর **annihilator**, আর লিখব $\text{ann}(S)$.

Example 9: Define annihilator of a subspace.[2] **(2021.CC12.3gp1)**

SOLUTION:

> DEFINITION: **Annihilator**
>
> Let V be a vector space with dual V^*. Let $S \subseteq V$ be any subset. Then by the annihilator of S we mean the following subset of V^*:
>
> $$\text{ann}(S) = \{f \in V^* \ : \ \forall \vec{v} \in S \quad f(\vec{v}) = 0\}.$$

লক্ষ্য করো, যদিও প্রশ্নে subspace বলেছিল, আমরা কিন্তু সংজ্ঞাটা দিয়েছি যেকোনো subset-এর জন্য। আসলে এখানে subspace-এর কোনো ভূমিকা নেই, যেকোনো subset-এর জন্যই একই সংজ্ঞা। যেহেতু subspace-রা subset-ও হয়, তাই ওই একই সংজ্ঞা subspace-দের বেলাতেও খাটে। ■

একটা উদাহরণ দেখা যাক।

Example 10: ধরো $V = \mathbb{R}^3$ আর $S = \{(1, 2, 0),\ (2, 3, 0)\}$. তবে $\text{ann}(S)$ বার করো।

SOLUTION: এখানে V^*-এর সদস্যদের দেখতে $f(x, y, z) = ax + by + cz$-এর মত। তার মানে আমরা সেইসব a, b, c খুঁজছি, যাদের জন্য $f(1, 2, 0) = 0$ এবং $f(2, 3, 0) = 0$ হয়, মানে

$$a + 2b = 2a + 3b = 0.$$

এর ফলে $a = b = 0$ হতে বাধ্য। কিন্তু c-এর উপর তো কোনো শর্ত নেই, তাই ওটা যা খুশি হতে পারে। তাই

$$\text{ann}(S) = \{f(x, y, z) = cz \ : \ c \in \mathbb{R}\}.$$

■

এখানে S-টা finite ছিল। এবার একটা উদাহরণ দেখি, যেখানে S-টা infinite.

Example 11: ধরো $\mathbb{R}^3$-এর মধ্যে S হল পুরো x-axis-টা। তবে $\text{ann}(S)$ কী হবে?

SOLUTION: এখানে এমন $f(x, y, z) = ax + by + cz$ খুঁজতে হবে, যাতে $\forall x \in \mathbb{R} \quad f(x, 0, 0) = 0$ হয়, মানে $ax = 0$ হয়। সেটা হবে একমাত্র $a = 0$ হলেই। ব্যস্, আর কোনো শর্ত নেই। তাই b আর c যা খুশি হতে পারে। সুতরাং--

$$\text{ann}(S) = \{f(x, y, z) = by + cz \ : \ b, c \in \mathbb{R}\}.$$

■

Exercise 10: নীচে প্রতিক্ষেত্রে বিভিন্ন $\mathbb{R}^n$-এর subset দিয়েছি। তোমার কাজ হল তাদের annihilator বার করা।

1. $\{\vec{0}\} \subseteq \mathbb{R}^n$.

2. পুরো $\mathbb{R}^n$.

3. $\mathbb{R}^3$-র মধ্যে x-axis আর y-axis-এর union, অর্থাৎ $\{(x, y, 0) \ : \ x = 0 \text{ or } y = 0\}$.

4. $\mathbb{R}^2$-এর মধ্যে unit circle, মানে $\{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 = 1\}$.

5. $\mathbb{R}^2$-এর মধ্যে open unit disc, মানে $\{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 < 1\}$.

6. $\phi \subseteq \mathbb{R}^n$.

■

Example 12: If $W = \{(x, y, z) \ : \ x - 2y - 3z = 0\}$ is a subspace of $\mathbb{R}^3$, find the annihilator of W.[3] **(2021.CC12.3gp2)**

SOLUTION: প্রথমে W-এর সদস্যদের general একটা রূপ লিখে নেওয়া যাক--

> A vector is in W if and only if it is of the form $(2y + 3z, y, z)$.

এটা কী করে এল বুঝলে তো? $\mathbb{R}^3$-এর যেকোনো সদস্যই (x, y, z) রূপের। সেটা W-এর মধ্যে থাকার শর্ত হল $x - 2y - 3z = 0$ হওয়া, মানে $x = 2y + 3z$ হওয়া। তাই (x, y, z)-এর মধ্যে x-এর জায়গায় $2y + 3z$ বসিয়ে দিয়েছি।

> So $\text{ann}(W)$ consists of all f of the form $f(x, y, z) = ax + by + cz$, where
>
> $$\forall y, z \in \mathbb{R} \quad a(2y + 3z) + by + cz = 0,$$
>
> ie,
>
> $$\forall y, z \in \mathbb{R} \quad (2a + b)y + (3a + c)z = 0.$$

এখানে y, z যা খুশি নেওয়া যাচ্ছে, সুতরাং একবার $(y, z) = (1, 0)$ এবং তারপর $(y, z) = (0, 1)$ নিলেই--

> Taking $y = 1, z = 0$ we have $2a + b = 0$, or $b = -2a$.
> Again taking $y = 0, z = 1$ we have $3a + c = 0$, ie, $c = -3a$.
> Hence
>
> $$\text{ann}(W) = \{f(x, y, z) = ax - 2ay - 3az \ : \ a \in \mathbb{R}\} = \text{span}\{x - 2y - 3z\}.$$

■

Exercise 11: If $W = \{(x, y, z) \ : \ 3x - 4y - 5z = 0\}$ is a subspace of $\mathbb{R}^3$, find the annihilator of W. ■

72.3 ann(S)-এর ধর্ম

এবার annihilator-দের কিছু ধর্ম আলোচনা করব। আমরা এ পর্যন্ত খালি এটুকুই জানি যে, $S \subseteq V$ হলে ann$(S) \subseteq V^*$ হয়। কিন্তু ann(S)-টা V^*-এর মধ্যে খালি একটা subset-ই নয়, এর কিছু বৈশিষ্ট্য আছে। যদি S-টা একটা subspace হয়, তবে S-এর dimension-এর সাথে ann(S)-এর dimension-এর কিছু সম্পর্ক আছে। এইসবই আলোচনা করব এবার। প্রথমে S-কে খালি একটা subset ধরে এগোব। পরে দেখব S একটা subspace হলে কী হয়।

72.3.1 যেখানে S যেকোনো subset

ann(S)-এর সবচেয়ে গুরুত্বপূর্ণ বৈশিষ্ট্য এই যে, সেটা সর্বদাই একটা subspace হতে বাধ্য, তা S-টা যাই subset হোক না কেন। কথাটা একটা theorem হিসেবে গুছিয়ে লিখে রাখি--

THEOREM

For any subset S of a vector space V, its annihilator ann(S) is a subspace of V^*.

Proof:

> Take any $S \subseteq V$.

এবার ann(S)-কে একটা subspace দেখানোর জন্য কী দেখাতে হবে? এক, এর মধ্যে zero vector আছে, আর দুই, এটা closed under linear combination.

> The zero vector in V^* is the linear functional that always takes the value 0. This is in ann(S).

কারণ এই zero linear functional-টা সবাইকেই শূন্যতে পাঠায়, তাই S-এর সদস্যদেরও রেহাই দেবে না।

> Also, if $\alpha, \beta \in \mathbb{R}$ and $f, g \in \text{ann}(S)$, then $\alpha f + \beta g \in \text{ann}(S)$.
>
> ⟦ Because:
>
> For any $\vec{v} \in S$ we have
>
> $$(\alpha f + \beta g)(\vec{v}) = \alpha f(\vec{v}) + \beta g(\vec{v}) = \alpha 0 + \beta 0 = 0.$$
>
> ⟧
>
> Thus ann(S) is a subspace, as required.

[Q.E.D]

এবার S-এর সঙ্গে ann(S)-এর একটা সম্পর্কের কথা বলি। ধরো $f \in V^*$ নিলাম। তবে সেটা কখন ann(S)-এর মধ্যে থাকবে? যখন S-এর প্রতিটা সদস্যকেই f শূন্যতে পাঠাবে। তার মানে S-এর প্রতিটা সদস্য f-এর উপর একটা করে নতুন শর্ত চাপাচ্ছে। তাই S যতই বড় হবে, শর্তের পরিমাণও ততই বাড়বে, ফলে S বড় হলে ann(S) কখনো বাড়তে পারে না। যেমন ধরো, একটা চাকরীর জন্য খালি হায়ার সেকণ্ডারী পাশ হলেই চলে, আর অন্য একটা চাকরীর জন্য হায়ার সেকণ্ডারী পাশের সঙ্গে আবার গীটার বাজাতেও জানতে হয়। তবে বুঝতেই পারছ, দ্বিতীয় চাকরীর জন্য যোগ্য লোকের সংখ্যা প্রথমটার চেয়ে বেশী হতে পারে না।

অবশ্য তা বলে ভেবে বোসো না যেন যে, S-কে বড় করলেই $\text{ann}(S)$ ছোটো হয়ে যাবে। যেমন ধরো $S = \{(1,0,0)\}$ ছিল, সেটাকে বাড়িয়ে $S = \{(1,0,0),(2,0,0)\}$ করলাম। তবে $\text{ann}(S)$ একই থাকবে, কারণ যেসব linear functional-এ $(1,0,0)$ কাটা পড়বে, সেসব linear functional-এই $(2,0,0) = 2 \times (1,0,0)$-ও কাটা পড়তে বাধ্য। একই যুক্তিতে এটাও বোঝা যায় যে, যদি S-এর মধ্যে ওর সদস্যদেরই কোনো নতুন linear combination ঢোকাই, তবে $\text{ann}(S)$ অপরিবর্তিত থাকবে। এ থেকেই প্রমাণ হয়ে যাবে নীচের theorem-টা।

THEOREM

Let V be a vector space. Let $S \subseteq V$ be some subset. Then $\text{ann}(\text{span}(S)) = \text{ann}(S)$.

Exercise 12: নীচের প্রতিটা বাক্য ঠিক না ভুল বলো। যদি ভুল হয়, তবে একটা করে counter-example দিও--

1. যদি $S \subseteq T$ হয়, তবে $\text{ann}(S) \subseteq \text{ann}(T)$ হবেই।
2. যদি $S \subseteq T$ হয়, তবে $\text{ann}(S) \supseteq \text{ann}(T)$ হবেই।
3. যদি $S \subsetneq T$ হয়, তবে $\text{ann}(S) \supsetneq \text{ann}(T)$ হবেই।

■

Example 13: Let W be the subspace of $\mathbb{R}^3$ spanned by $(1,1,0)$ and $(0,1,1)$. Find a basis of the annihilator of W.[5] **(2020.CC12.3c)**

SOLUTION:

> To find a basis for $\text{ann}(\text{span}(\{(1,1,0),\ (0,1,1)\})) = \text{ann}(\{(1,1,0),\ (0,1,1)\})$.

এখানেই আমরা আগের theorem-টা লাগালাম। এতে পুরো span-টার সবাইকে আলাদা করে ঘায়েল করার চেষ্টা না করে খালি দুটো vector-কে ঘায়েল করতে পারলেই হবে।

> Thus, for $f(x,y,z) = ax + by + cz$, we have $f \in \text{ann}(W)$ if and only if
>
> $$f(1,1,0) = 0 \text{ and } f(0,1,1) = 0,$$
>
> ie, $a + b = 0$ and $b + c = 0$, or
>
> $$a = c = -b.$$
>
> Hence $\text{ann}(W) = \{f(x,y,z) = -bx + by - bz \ : \ b \in \mathbb{R}\}$.

তার মানে $\text{ann}(W)$-র মধ্যে রয়েছে কেবল $-x+y-z$-এর সব multiple-রাই। বোঝাই যাচ্ছে যে, $\{-x+y-z\}$-ই হবে $\text{ann}(W)$-এর একটা basis. সেটাই গুছিয়ে লিখে দেব।

> Let $g(x,y,z) = -x + y - z$. Then $g(x,y,z)$ spans $\text{ann}(W)$.
> Also $g(x,y,z)$ is not the zero linear functional.
> $\because \{g(x,y,z)\}$ is also linearly independent.

> Thus, a required basis is $\{g(x,y,z) = -x+y-z\}$.

■

Exercise 13: Let W be a subspace of $\mathbb{R}^4$ spanned by $(1,2,-3,4)$, $(1,3,-2,6)$ and $(1,4,-1,8)$. Find a basis of the annihilator of W.[5] **(2022.CC12.3e)**
HINT: মাঝখানের vector-টা কিন্তু বাকিদুটোর গড়, তাই ওদের span-এর মধ্যে আছে। অতএব ওটাকে নিয়ে আলাদা করে মাথা না ঘামালেও হবে।
■

72.3.2 যেখানে S একটা subspace

যদি S একটা subspace হয়, তবে ann(S)-এর dimension আর S-এর dimension-এর মধ্যে কিছু সম্পর্ক আছে। আমরা এই বইতে আগে subspace-দের orthogonal complement নিয়ে আলোচনা করেছিলাম। Annihilator-এর সঙ্গে orthogonal complement-এর বেশ সাদৃশ্য আছে। অবশ্য orthogonal complement-এর ধারণাটা কেবল inner product space-দের বেলাতেই খাটে (তার জন্য field-টাকে $\mathbb{R}$ বা $\mathbb{C}$ নিতে হয়), যদিও annihilator-এর ধারণাটা যেকোনো field-এর উপরে যেকোনো vector space-এর ক্ষেত্রেই প্রযোজ্য।

> THEOREM
>
> If S is a subspace of a finite dimensional vector space, V, then
>
> $$\dim(\text{ann}(S)) = \dim(V) - \dim(S).$$

Proof: Orthogonal complement-এর সঙ্গে মিলটা চোখে পড়ছে তো? V যদি একটা inner product space হত, আর তার মধ্যে S হত একটা subspace , তবে $\dim(S^\perp) = \dim(V) - \dim(S)$ হত। সেই প্রমাণটা যদি মনে থেকে থাকে, তবে নীচের প্রমাণের ধাপগুলোও পরিচিত ঠেকবে।

> Let $\dim(V) = n$ and $\dim(S) = m \leq n$.

দেখাতে চাইছি $\dim(\text{ann}(S)) = n-m$. তার জন্য আমরা হাতেনাতে ann(S)-এর একটা basis দেখিয়ে দেব, যার সাইজ $n-m$.

> Let $\{\vec{v}_1, ..., \vec{v}_m\}$ be a basis of S.
> We extend it to a basis $\{\vec{v}_1, ..., \vec{v}_m, \vec{v}_{m+1}, ..., \vec{v}_n\}$ of V.

এই extension-এর ফলে বাড়তি $n-m$-খানা vector যোগ হল। এরা মিলে কি ann(S)-র একটা basis হয়? না, ওরা তো V-এর মধ্যে রয়েছে, আমাদের তো যেতে হবে V^*-এ, কারণ $\text{ann}(S) \subseteq V^*$. কী করে $\vec{v}_i$-দের থেকে V^*-এর সদস্য বানানো যায়? উত্তর হল, dual basis নিলে--

> Let $\{f_1, ..., f_n\}$ be its dual basis.

এবার দেখাব যে, $\{f_{m+1}, ..., f_n\}$ হবে ann(S)-এর একটা basis. তার জন্য কী কী দেখাতে হবে? প্রথমতঃ দেখাতে হবে যে, এরা সবাই ann(S)-এর মধ্যে আছে--

> Then $\forall i \in \{m+1, ..., n\} \quad f_i \in \text{ann}(S)$.
>
> ⟦ <u>Because:</u>
>
> Take any $\vec{s} \in S$. Then $\vec{s} = b_1\vec{v}_1 + \cdots + b_m\vec{v}_m$ for some scalars $b_1, ..., b_m$.
> Hence for each $i \in \{m+1, ..., n\}$, we have
>
> $$f(\vec{s}) = f(b_1\vec{v}_1 + \cdots + b_m\vec{v}_m) = b_1 f_i(\vec{v}_1) + \cdots + b_m f_i(\vec{v}_m) = 0,$$
>
> since $f_i(\vec{v}_j) = 0$ when $i \neq j$.
>
> ⟧

তারপর দেখাতে হবে spanning.

> Next, we shall show that $\{f_{m+1}, ..., f_n\}$ spans $\text{ann}(S)$.

তার মানে যেকোনো $f \in \text{ann}(S)$ নিয়ে দেখাব যে, সেটাকে $f_{m+1}, \ldots, f_n$-এর linear combination হিসেবে লেখা যায়।

> Take any $f \in \text{ann}(S)$.
>
> Let $b_i = f(\vec{v}_i)$ for $i = m+1, ..., n$. Then $f = b_{m+1}f_{m+1} + \cdots + b_n f_n$.

এটা দেখানোয় বেশ একটু কৌশল আছে--

> ⟦ <u>Because:</u>
>
> Let $g = b_{m+1}f_{m+1} + \cdots + b_n f_n$.
> Then, for $i = 1, ..., m$, we have $f(\vec{v}_i) = 0 = g(\vec{v}_i)$.
> Also, for $i = m+1, ..., n$, we have $f(\vec{v}_i) = b_i = g(\vec{v}_i)$.
> Thus f and g are two linear functionals matching over a basis. So they must be identical.
>
> ⟧

বাকি রইল খালি linearly independent দেখানো--

> The set $\{f_{m+1}, ..., f_n\}$ is linearly independent.
>
> ⟦ <u>Because:</u>
>
> Let $c_{m+1}f_{m+1} + \cdots + c_n f_n =$ the zero functional.
> Then for each $i = m+1, ..., n$, we have
>
> $$(c_{m+1}f_{m+1} + \cdots + c_n f_n)(\vec{v}_i) = 0,$$
>
> or $c_i = 0$, since $f_i(\vec{v}_j) = 0$ if $i \neq j$, and $f_i(\vec{v}_j) = 1$ if $i = j$.
>
> ⟧

This completes the proof.

[Q.E.D]

এবার আরেকটা theorem, যেখানেও annihilator-দের সঙ্গে orthogonal complement-দের সাদৃশ্যটা চোখে পড়ার মত।

> **THEOREM**
> If S, T are two subspaces of a vector space, then
> $$\text{ann}(S+T) = \text{ann}(S) \cap \text{ann}(T).$$

Proof:

Step 1: $\text{ann}(S+T) \subseteq \text{ann}(S) \cap \text{ann}(T)$:

Take any $f \in \text{ann}(S+T)$.

Then $\forall \vec{v} \in S+T \quad f(\vec{v}) = 0$.

Since $S, T \subseteq S+T$, hence $\forall \vec{v} \in S \quad f(\vec{v}) = 0$ and $\forall \vec{v} \in T \quad f(\vec{v}) = 0$.

So $f \in \text{ann}(S)$ and $f \in \text{ann}(T)$. Thus, $f \in \text{ann}(S) \cap \text{ann}(T)$, as required.

Step 2: $\text{ann}(S+T) \supseteq \text{ann}(S) \cap \text{ann}(T)$:

Take any $f \in \text{ann}(S) \cap \text{ann}(T)$.

Then $\forall \vec{v} \in S \quad f(\vec{v}) = 0$ and $\forall \vec{v} \in T \quad f(\vec{v}) = 0$.

Take any $\vec{v} \in S+T$. Then $\vec{v} = \vec{s} + \vec{t}$ for some $\vec{s} \in S$ and some $\vec{t} \in T$.

So $f(\vec{v}) = f(\vec{s} + \vec{t}) = f(\vec{s}) + f(\vec{t}) = 0$.

Hence $f \in \text{ann}(S+T)$, as required.

[Q.E.D]

> **THEOREM**
> If S, T are two subspaces of a vector space, then
> $$\text{ann}(S \cap T) \supseteq \text{ann}(S) + \text{ann}(T).$$
> If the vector space is finite dimensional, then we have
> $$\text{ann}(S \cap T) = \text{ann}(S) + \text{ann}(T).$$

Proof:

> First part:
>
> Take any $f \in \mathrm{ann}(S) + \mathrm{ann}(T)$. Then $f = g + h$ for some $g \in \mathrm{ann}(S)$ and $h \in \mathrm{ann}(T)$.

দেখাব যে, $f \in \mathrm{ann}(S \cap T)$, মানে যাই $\vec{v} \in \cap T$ নাও না কেন, $f(\vec{v}) = 0$ হবেই।

> Take any $\vec{v} \in S \cap T$.
>
> Then $f(\vec{v}) = g(\vec{v}) + h(\vec{v}) = 0$.
>
> Hence $f \in \mathrm{ann}(S \cap T)$. Thus, $\mathrm{ann}(S \cap T) \supseteq \mathrm{ann}(S) + \mathrm{ann}(T)$, as required.

এবার দ্বিতীয় অংশ। এখানে দেখাব যে $\mathrm{ann}(S + T)$ আর $\mathrm{ann}(S) + \mathrm{ann}(T)$-এর dimension সমান। এর জন্য এই তথ্যটা লাগবে--যদি U, V দুটো subspace হয়, তবে $\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)$.

> Second part: Let the finite dimensional vector space have dimension n.
>
> Also, let $s = \dim(S)$, and $t = \dim(T)$.
>
> Then
>
> $$\begin{aligned} & \dim(\mathrm{ann}(S) + \mathrm{ann}(T)) \\ = \ & \dim(\mathrm{ann}(S)) + \dim(\mathrm{ann}(T)) - \dim(\mathrm{ann}(S) \cap \mathrm{ann}(T)) \\ = \ & \dim(\mathrm{ann}(S)) + \dim(\mathrm{ann}(T)) - \dim(\mathrm{ann}(S + T)) \\ = \ & [n - \dim(S)] + [n - \dim(T)] - [n - \dim(S + T)] \quad [\text{ by the last theorem}] \\ = \ & n - \dim(S) - \dim(T) + \dim(S) + \dim(T) - \dim(S \cap T) \\ = \ & n - \dim(S \cap T) \\ = \ & \dim(\mathrm{ann}(S \cap T)) \quad [\text{ by the last theorem}]. \end{aligned}$$
>
> Since we had $\mathrm{ann}(S \cap T) \supseteq \mathrm{ann}(S) + \mathrm{ann}(T)$, this proves the required equality.

[Q.E.D]

> THEOREM
>
> If T is a direct complement of S in a finite dimensional vector space, V, then $\mathrm{ann}(T)$ is a direct complement of $\mathrm{ann}(S)$ in V^*.

Proof: আগের theorem দুটো লাগালে অতি সংক্ষেপেই প্রমাণটা হয়ে যাবে। মানে রেখো "T হল S-এর একটা direct complement." মানে হল "$S + T = V$ এবং $S \cap T = \{\vec{0}\}$." গুছিয়ে লেখার দায়িত্ব তোমাকেই দিলাম। **[Q.E.D]**

Answers

1. আমরা এখানে linear functional খুঁজছি। যেকোনো linear transformation-র বেলাতেই আমরা জানি যে, একটা basis-এর উপর তাদের value বেঁধে দিলেই linear transformation-টা নির্ধারিত হয়ে যায়। এখানে আমরা

প্রতিটা f_i-এর value প্রতিটা $\vec{v}_j$-র জন্য বেঁধে দিয়েছি। **2.** Linearly independent দেখানোর জন্য আমরা $a_1f_1 + \cdots + a_nf_n = 0$ নিয়েছিলাম। এর দুদিকেই linear functional আছে। যদি $\vec{v}_i$-এর উপরে লাগাই, তবে বাঁদিকে পাবে $a_1f_1(\vec{v}_i) + \cdots + a_nf_n(\vec{v}_i)$, আর ডান দিকে পাবে 0. আমরা জানি $f_j(\vec{v}_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$. সুতরাং পাচ্ছি $a_i = 0$.

Spanning দেখানোর জন্য আমরা $a_i = f(\vec{v}_i)$ নিয়েছিলাম। সুতরাং $(a_1f_1 + \cdots + a_nf_n)(x) = \sum f(\vec{v}_i)f_i(\vec{x})$. আমরা দেখাতে চাই যে, এটা আসলে $f(\vec{x})$. যেহেতু আমরা linear transformation নিয়ে কাজ করছি, তাই যাবতীয় $\vec{x}$-এর জন্য আলাদা করে না দেখিয়ে খালি কোনো একটা basis-এর জন্য দেখালেই চলবে। যদি $\vec{x}$-এর জায়গায় $\vec{v}_j$ বসাই তবে $\sum f(\vec{v}_i)f_i(\vec{v}_j) = f(\vec{v}_j)$, কারণ $i \neq j$-এর বেলায় $f_i(\vec{v}_j) = 0$. সুতরাং $\{\vec{v}_1, ..., \vec{v}_n\}$-এর উপরে $\sum f(\vec{v}_i)f_i$ আর f মিলে যাচ্ছে। তাই $\sum f(\vec{v}_i)f_i \equiv f$, ঠিক যেটা দেখানো দরকার ছিল। **3.** (B). যুক্তি হল $f_i(\vec{v}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$ দেখিয়ে দেওয়া। যেহেতু dual basis-টা unique, তাই (B) যেই ঠিক দেখানো হয়ে গেল, বাকিগুলো আর ঠিক হতে পারে না। **4.** $\left\{\left(\frac{2}{5}, -\frac{3}{10}, -\frac{1}{10}\right), \left(\frac{3}{5}, \frac{3}{10}, \frac{1}{10}\right), \left(\frac{1}{5}, \frac{1}{10}, -\frac{3}{10}\right)\right\}$ **5.** একটা ধাপ করে দেখাচ্ছি। আমাদের দেওয়া আছে $f_1(a_1, a_2, a_3) = 1$, মানে $p_1a_1 + p_2a_2 + p_3a_3 = 1$. দেখাতে হবে $g_1(p_1, p_2, p_3) = 1$ মানে $a_1p_1 + a_2p_2 + a_3p_3 = 1$. কিন্তু সেটাই তো দেওয়া ছিল!

6. $\left\{\left(-\frac{16}{9}, \frac{14}{9}, -\frac{1}{9}\right), \left(\frac{8}{9}, -\frac{7}{9}, \frac{2}{9}\right), \left(-\frac{1}{9}, \frac{2}{9}, -\frac{1}{9}\right)\right\}$. **7.** প্রতিটা b_i-কেই আলাদা করে শূন্য দেখানোর জন্য

$$v(x) = (x - a_1) \cdots (x - a_{i-1})(x - a_{i+1}) \cdots (x - a_n)$$

বসাও। তবে $e_{a_i}(v) \neq 0$, এবং $j \neq i$ হলে $e_{a_j}(v) = 0$. সুতরাং $\sum_j b_je_{a_j}(v) = b_ie_{a_i}(v) = 0$, এবং তা থেকে $b_i = 0$ হয়ে যাচ্ছে।

$V^* \not\cong V$ দেখানোর জন্য লক্ষ্য করো V-এর একটা basis হল $\{1, x, x^2, ...\}$, যেটা countable. সুতরাং V-এর সঙ্গে isomorphic কোনো vector space-এর মধ্যে uncountable একটা linearly independent set থাকতে পারে না।

8. $(1, 2, 3, 4) = 21\vec{v}_1 + 19\vec{v}_2 + 12\vec{v}_3 + 17\vec{v}_4$. **9.** $4x + 0y$. **10.** **(1)** $\mathbb{R}^n$-এর উপরে যাবতীয় linear functional-এর set. **(2)** $\mathbb{R}^n$-এর উপরে কেবলমাত্র zero linear functional দিয়ে তৈরী singleton set. **(3)** $\mathbb{R}^3$-এর উপরে $f(x, y, z) = cz$ চেহারার সব linear functional-দের set. **(4)** $\mathbb{R}^2$-এর উপরে কেবলমাত্র zero linear functional দিয়ে তৈরী singleton set. **(5)** $\mathbb{R}^2$-এর উপরে কেবলমাত্র zero linear functional দিয়ে তৈরী singleton set. **(6)** $\mathbb{R}^n$-এর উপরে যাবতীয় linear functional-এর set.

11. $\text{ann}(W) = \text{span}\{3x - 4y - 5z\}$. **12.** **(1)** ভুল। $\mathbb{R}^3$-র মধ্যে $S = x$-axis আর $T = xy$-plane নিলে $\text{ann}(S) = \text{span}\{y, z\}$ এবং $\text{ann}(T) = \text{span}\{z\}$. দেখাই যাচ্ছে $\text{ann}(S) \not\subseteq \text{ann}(T)$. **(2)** ঠিক। **(3)** ভুল। $\mathbb{R}$-এর মধ্যে $S = [0, 1]$ আর $T = [0, 2]$ নিলে $\text{ann}(S) = \text{ann}(T) = \{\text{zero linear functional}\}$ হয়, যদিও $S \subsetneq T$.

13. $\{5x - y + z,\ 2y - w\}$.

Index